

# Psychological Review

THEODORE M. NEWCOMB, Editor  
UNIVERSITY OF MICHIGAN  
Lorrairie Bouthilet, Managing Editor

---

## CONTENTS

Surface and Depth in the Individual Personality .....	NEVITT SANFORD 349
Consciousness as a Regulatory Field: A Theory of Psychopathology .....	REX M. COLLIER 360
Frustration Theory: Restatement and Extension .....	NORMAN R. F. MAIER 370
The Inadequacy of the Hullian Derivations of Reasoning and Latent Learning .....	J. A. DEUTSCH 389
Perception: Toward the Recovery of a Definition .....	JULIAN HOCHBERG 400
Secondary Motivation Through Classical Conditioning: A Reconsideration of the Nature of Backward Conditioning .....	JOHN A. BARLOW 406

---

This is the last issue of Volume 63.  
Volume contents and title page ap-  
pear herein.

---

PUBLISHED SIMONTHLY BY THE  
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

## CONSULTING EDITORS

SOLOMON E. ASCH  
ROBERT R. BLAKE  
STUART W. COOK  
CLYDE H. COOMBS  
LEON FESTINGER  
W. R. GARNER  
JAMES J. GIBSON  
D. O. HEBB  
HARRY HELSON  
E. R. HILGARD  
CARL I. HOVLAND  
E. LOWELL KELLY  
DAVID KRECH  
ROBERT W. LEEPER  
KENNETH MACCORQUODALE

ROBERT B. MACLEOD  
DAVID C. MCCLELLAND  
GEORGE A. MILLER  
GARDNER MURPHY  
OSCAR OESER  
CARL PFAFFMANN  
CARROLL C. PRATT  
JOHN P. SEWARD  
DAVID SHAKOW  
RICHARD L. SOLOMON  
ELIOT STELLAR  
S. S. STEVENS  
ERIC TRIST  
EDWARD L. WALKER  
ROBERT W. WHITE

The *Psychological Review* is devoted to theoretical articles of significance to any area of psychology. Except for occasional articles solicited by the Editor, manuscripts exceeding twelve printed pages (about 7,500 words) are not accepted. Ordinarily manuscripts which consist primarily of original reports of research should be submitted to other journals.

Because of the large number of manuscripts submitted, there is an inevitable publication lag of several months. Authors may avoid this delay if they are prepared to pay the costs of publishing their own articles; the appearance of articles by other contributors is not thereby delayed.

Tables, footnotes, and references should appear on separate pages; all of these, as well as the text, should be typed double-spaced throughout, in all manuscripts submitted. All manuscripts should be submitted in duplicate. Original figures are prepared for publication; duplicate figures may be photographic or pencil-drawn copies. Authors are cautioned to retain a copy of the manuscript to guard against loss in the mail. Manuscripts should be addressed to the Editor, Dr. Theodore M. Newcomb, Doctoral Program in Social Psychology, University of Michigan, Ann Arbor, Michigan.

PUBLISHED BIMONTHLY BY THE  
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.  
PRINCE AND LEMON STREETS, LANCASTER, PA.  
1333 SIXTEENTH ST. N. W., WASHINGTON 6, D. C.

\$3.00 volume

\$1.50 issue

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879.

Acceptance for mailing at the special rate of postage provided for in paragraph (6-3), Section 3440, P. L. & E. of 1946, authorized Jan. 2, 1948.

Send all communications, including address changes, to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must arrive by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1956 by the American Psychological Association, Inc.

# THE PSYCHOLOGICAL REVIEW

## SURFACE AND DEPTH IN THE INDIVIDUAL PERSONALITY<sup>1</sup>

NEVITT SANFORD

*Department of Psychology, University of California, Berkeley, and Mary Conover  
Mellon Foundation, Vassar College*

In his undergraduate course in personality, the writer has for some time followed the practice of beginning with the "surface"—with the descriptive, the objectively observable, the phenotypical—and proceeding inward, as it were, to the genotypical, the underlying dynamics, the "depths." He has been forced to admit, however, that his conceptions of surface and of depth have been vague, and he has wondered whether his colleagues have achieved greater clarity.

A quick survey shows that psychologists have been much taken up with problems of surface and depth. The literature abounds with terms such as those I have just used—the descriptive vs. the inferred, phenotypical-genotypical, the central and the peripheral, the superficial and the deep. Psychologists, like the literary characterologists before them, have usually conceived of personality as a whole, existing in time and space, and having dimensionality; and to suggest this, they have used analogies from geology, oceanography, architecture, horticulture, and anthropology, as well as geometry and topology. Thus we find references to the surface, the superstructure, the shell or skin or peel,

the mask or veil, the fringe, the periphery as opposed to the subterranean, the submerged, the foundations or underpinnings, the kernel or core or heart, the hidden or mysterious, the center, the central.

Surface and depth, though common in the literature, are nowhere used as technical terms. Depth psychology, however, is frequently used to refer to psychoanalysis. The terms usually have evaluative connotations. Everyone wants to be deep or to go deep; no one wants to be shallow or superficial. On the other hand, there seems to be some ambivalence in the matter, for there is the danger that if, like the psychoanalysts, one goes down deep, he might come up dirty.

As evidence of the seriousness with which the matter of surface and depth is viewed in contemporary thought I may quote, not from a psychologist, but from the biologist, Bertalanffy. He distinguishes between the "depth personality with its primeval instincts, emotions and appetites" and "the day personality of consciousness," and goes on to say that "the antagonism between the levels of personality is at the bottom of the human tragedy . . . whether the levels of personality can be properly adjusted is the question upon which man's future depends" (4, p. 37).

What are the levels? How does one

<sup>1</sup> A shortened version of a paper presented as the Presidential address of the Division of Personality and Social Psychology of the American Psychological Association, San Francisco, September 2, 1955.

distinguish one from another? And how do the levels interact, if indeed they do?

In taking a brief glance at the history of these problems in psychology, we may start with Freud's division of the mind into conscious, preconscious, unconscious. It is interesting to note that this topography, though it was soon to be replaced in the forefront of psychoanalytic thinking by the trichotomy of Superego, Ego, Id, has been assumed by psychoanalytic writers and has never been altered. It has not been seriously challenged by neo-Freudians. Jung, however, did challenge it, as early as 1912, and by adding the concepts of persona and collective unconscious, he presented a topography of more layers or levels than the original. Both of these schemes were very much in the picture when psychoanalysis began to make an impression in the Universities.

Murray, whose *Explorations in Personality* (14) was published in 1938, distinguished between manifest (overt, objectified) and latent (covert, subjectified) needs, and sought to obtain independent measures of the two. It was with the latter, particularly, in mind that Morgan and Murray (13) invented the Thematic Apperception Test (TAT).

Some of the early work with the TAT, in which the attempt was made to relate variables in the subjects' story productions to variables of overt behavior (15), stands among the first efforts at objective and quantitative study of variables conceived to be operating on different levels. But this work was fairly naive, as it has turned out. Today it would be fairly generally agreed that scores based on all the manifestations of needs in stories cannot be regarded as measures either of fantasy or of behavior. The point, made most clearly by Tomkins (17), is that TAT contents may be indices of behavior, of conscious fantasy or of unconscious motives, and that the problem is to discover which indices are

to be referred to which level. This way of looking at things was adopted and furthered by Betty Aron in her manual on the TAT (3).

Frenkel-Brunswick (8), beginning in 1940, has consistently stressed the discrepancies between surface behavior and underlying motives, and has underlined the enormous complexity of the relationships between these two levels of functioning. In a 1942 monograph (9) she educed some evidences of alternative surface manifestations of the same motivating force.

When it came to the work on *The Authoritarian Personality* (1), the authors took it for granted that they should have to distinguish the opinions and attitudes that the subject would express publicly from those he would express only to an intimate friend, and that both these groups of phenomena would have to be distinguished from dispositions that the subject could not admit to himself.

This way of thinking about levels has been developed most fully, perhaps, by Freedman, Leary, Ossorio, and Coffey (7), who distinguish the public, the conscious, and the private levels of personality, and offer operational definitions of variables on each level, in terms of a theory according to which one kind of variable might be related to all the others.

The line of development just indicated, which has been close to the clinic and close to psychoanalysis, is of course but a part of the recent history of thought and research bearing on the general problem of surface and depth. An adequate account would give much attention to the early work of Kurt Lewin, which made phenotype, genotype, peripheral, central, layer, strata, region, household words in psychology; to the writings of Gordon Allport (2), whose trait psychology makes large use of the distinctions phenotypical vs. gen-



atypical, instrumental vs. motivational, specific vs. general or central or cardinal, pseudo vs. genuine, and whose compelling arguments for functional autonomy were at the same time, but unintentionally, arguments that there were other, presumably deeper forces that were not functionally autonomous; and to the work of Cattell (5) who showed that factorial methods could be applied to problems of dynamics as well as to the task of classifying behavior.

One result of all the history to which I am alluding—history which in very considerable part has as its theme what to do about, or to, or with psychoanalysis—is that we now have a number of concepts—at least nine, I believe—which have to be considered in a discussion of surface and depth. In contemporary writing, the surface-depth dimension is spoken of as if it were analogous to, or at least an attribute of, each of these nine concepts.

Since the concepts are very familiar, I need do little more than mention them here.

1. *Availability to consciousness.* The unconscious is almost always regarded as deeper than the preconscious and the conscious. Preconscious processes, presumably, may vary among themselves with respect to their availability to consciousness.

2. *Availability to the motoric.* The inhibited is usually regarded as deeper than that which is freely expressed in action.

3. *Place in a developmental sequence.* In other schools of psychology besides psychoanalysis, that which is laid down, or set up, early in the individual's development, is commonly regarded as deeper than that which is established later.

4. *Biological as opposed to learned.* Biological needs, if they are admitted at all, usually are regarded as more basic, i.e., deeper, than derived needs; some-

thing more or less biological has a place at the bottom of most motivational pyramids.

5. *Neurologically "lower" or "higher."* Bertalanffy, in the paper quoted earlier (4), speaks of the cortex, "the organ of the day personality," as being "on top" of the palécephalon, "the organ of the depth personality." There is a certain awkwardness here, for what is neurologically more peripheral sometimes has to be regarded as deeper than that which is neurologically central.

6. *The genotypical vs. the phenotypical.* It is common to speak of the genotypical, the integrating, or determining, or ruling, or motivating as deeper than the phenotypical, the integrated, or determined, or ruled, or instrumental. It is also common to speak of the former as central and the latter as peripheral, though it is doubtful that this is justified if one sticks to Lewin's topology.

7. *Outer vs. inner.* These terms, of course, are also from topology, but not all writers have a topological model in mind when they speak of those features of the personality which are dependent upon immediate field conditions versus those which are characteristic of the individual in a deeper sense.

8. *Resistance to change.* The fixed or deep-seated is likely to be regarded as deep in a bad sense; the enduring or persistent as deep in a good sense.

9. *Openness to observation vs. the hidden or mysterious.* Since these conditions depend as much upon the observer as upon the subject, they should probably be left out of consideration.

It appears that the time is ripe for a serious new effort at clarification in this whole area, and to this end I should like to make a few suggestions.

The first suggestion is that we distinguish between concepts of structure and concepts of dynamics. I mean here structure in the formal rather than in the material sense. Personality can be

—and should be first of all—conceived as a whole, having parts which stand in various purely structural relations one to another. We can fill in the several cells or regions with such contents—needs, sentiments, or object cathexes, habits, cognitive dispositions—as we find appropriate, and then proceed to consider dynamics, the ways in which these contents interact, the mechanisms by which effects are achieved.

The next suggestion is that central-peripheral and inner-outer distinctions be adopted as concepts referring to structure, and that we hold ourselves strictly to this usage.

*Central-peripheral* would be used in the topological sense, as offered by Kurt Lewin in his later writings (11). *Central*, in this sense, means influencing and being influenced by relatively many other parts or regions of the person—and *peripheral* means the opposite. An aroused need or an instigation to goal activity would be central, while various activities brought into its service would be peripheral. This seems to be a very common instance of a process being central in the sense of influencing; probably it is a little awkward for most of us to think of the same process as being at the same time influenced by various other processes. But it might be profitable to think in that way. For example, a need to raise self-esteem, when it is aroused, will certainly mobilize in its service various action patterns—and thus be influencing; at the same time it is interesting to consider the sensitivity of such a need—how, in its subjective aspect, at least, everything that happens, over a wide area, seems to make a difference to it.

*Outer*, as opposed to *inner*, in the purely structural sense, refers of course to that which has a boundary in common with the boundary of the whole. Lewin used this conception of outer-inner, and went on to point out how,

in topology, a cell might be outer and yet central, inner and yet peripheral. He did not, however, offer many psychological examples of these states of affairs, nor did he or his students, as far as I know, make much use of these conceptions in considering the structure of personality.

By *outer* I shall mean here simply those layers or regions of the person that are involved in transactions with the external environment—by *inner*, those which are not. The conception would apply, quite literally, I think, to the case of a man of whom we say that he is so taken up with external matters that he has no time to be himself. Of ourselves we say that we are so involved with the urgent that we have no time for the important.

According to Lewin's topological theory, inner layers of the person are separated from the environment by surrounding outer layers, so that a process in an inner region cannot reach the outside save by passing through these outer layers. This theory does not seem to fit the facts of observation. Inner needs do not have to pass through outer regions in order to be expressed in behavior. They will be expressed directly, unless they are anchored inside—as in the case of a man who simply prefers to keep his best thoughts to himself—or unless the channels for their expression are blocked, either because these channels are filled with the traffic of day-to-day urgency or because a barrier has been erected against the specific need in question. Consider the case of a man whose everyday behavior is more or less dominated by a persona. We need not suppose that his inner needs can find expression only by passing through this outer layer, though no doubt they may influence what goes on there; instead, there are occasions on which the mask falls away and aspects of the inner personality stand directly revealed.

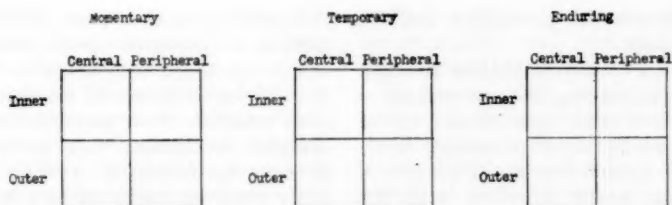


FIG. 1. Personality structures.

Now, since *central-peripheral* and *inner-outer* are independent aspects of structure, we have to consider that both inner and outer processes (contents) may be either central or peripheral. An interesting exercise is to draw a  $2 \times 2$  table, with *central-peripheral* along one axis and *inner-outer* along the other, and to set about filling in the squares with psychological contents.

But first let us remember that *time* is another factor that has to be taken into account. Personality structures may be momentary, temporary or enduring. So we need three of these  $2 \times 2$  tables.

Now I suggest that we have a scheme, formal and abstract as it is, that is inclusive with respect to the phenomena of personality. We are in a position to take the first step toward describing, in the terms of the same scheme, the phenomena of child development, of neurosis and psychosis, of normal adulthood.

Turning to the task of filling in the squares, let us limit ourselves, in the few minutes available now, to the trickier ones, or to those which upset, a little, our usual habits of thought. What is something that is inner, peripheral, and enduring? I suggest a tendency that was long ago more or less effectively insulated by repression, and in time has been more or less washed out as other dynamic organizations have ebbed and flowed, but which nevertheless remains as a potential, capable of becoming central when just the right stimulus comes

along. How about outer-central-enduring? The consistent instrumentalities, of taking a social role whose requirements are diverse and taxing, and which does not fit the individual very well, so that his inner needs find relatively little opportunity for expression, offer an example.

Inner-central-temporary would be represented, rather well I think, by the case of an individual involved in a developmental crisis. When, for example, an adolescent is taken up with the problem of independence, we find a structure in which very many needs, sentiments, abilities, cognitive readinesses are influenced by, even as they influence, this central core of striving, while other inner needs, even those which will be determining with respect to life goals, recede temporarily into the periphery.

How, according to this scheme, shall we describe hebephrenia? This is the extreme case of outer and peripheral, whether temporary or enduring. The individual is at the beck and call of the stimuli that chance to impinge upon him, and the connections among his responses are minimal. This, of course, corresponds to the condition of the infant who has not yet established any devices for the delay of response. Paranoia, on the other hand, might be thought of as the limiting case of centralism, everything being related, however unreasonably, to everything else, perhaps in a final desperate effort to prevent the dissolution of the bounda-

ries between the personality and the environment.

It seems to me possible that we might achieve, before long, fairly general agreement about some such scheme for the formal description of personality structure. I assume that the conditions of structure, to be described with this scheme, call for explanation by psychodynamic hypotheses.

In psychodynamics it is common to distinguish between adaptive and defensive mechanisms. When we say, as we commonly do, that similar behavior may have different determinants, the major classes of determination that we have in mind are (a) determination by conscious, functionally autonomous, ego-integrated motives, and (b) determination by motives that are defensive and unconscious. Another interesting exercise would be to go back to the tables shown in Fig. 1, and ask ourselves whether processes belonging to each square might not be either conscious or unconscious. It seems that they can be, and this is true whether we define "unconscious" merely as absence of awareness, or in the Freudian sense.

It is clear that we cannot use the terms *central* or *inner* to help distinguish our unconscious, defensive motives. As a matter of fact, it seems that, by and large, the greatest centrality is achieved in conscious functioning, when the use of symbols is brought to the highest level of development; and an inner need that is unconscious is not necessarily more inner than one that is conscious.

Now the important point: unconscious motives are not to be distinguished by calling them deep, in the sense of underneath—i.e., underneath strata which they must pass through in order to reach the surface. There are too many people who would like to get rid of unconscious motives by burying them in this way. The fact is that making a motive (or

some of its integrated effects, instrumentalities, or cognitive content) unconscious may be the very device by which that motive is permitted expression in overt behavior. More than this, whereas defensive mechanisms may operate to prevent environmental stimuli from being perceived, they may be powerless to prevent unconscious discrimination, whereby the most primitive, emotional, irrational needs of the persons come under the direct influence of the environment.

What I am saying here is that a major function of unconsciousness is that of keeping things separate, of preventing communication among parts which, if the person were functioning adaptively, would belong together.

These considerations are of crucial importance for social psychology. The social psychologist cannot assume that because he deals with the average, normal adult he need concern himself only with responses which are under the sway of the higher mental processes. On the contrary, social stimuli, such as propaganda, may make their appeal directly to unconscious forces; indeed, the technique of propaganda consists most essentially in finding ways to bypass the higher mental processes in reaching the responsive primitive needs. Similarly, the truly dangerous social or group phenomena, such as riots or totalitarian movements, and the truly intractable ones, such as social structures in institutions and organizations, usually involve the operation of unconscious mechanisms; in the former case such mechanisms have the crucial function of helping to remove ordinary restraints, while in the latter case the social structure may resist change just because it serves the defensive purposes of various individual members of the group. The social psychologist can ill afford not to be conversant with unconscious mental functioning.

Now let us consider a dynamic formulation of the kind that is frequently made in clinical cases. A case from *The Authoritarian Personality* (1) will do for the present purpose. Here the formulation was that this man's self-assertive behavior was an attempt to overcompensate for an underlying fear of weakness, which in its turn was based upon unconscious passive, homosexual impulses; these impulses could be understood as responses to, attempts to deal with, fear of the father, i.e., castration anxiety—a fear kept alive by aggression against the father, and which, in fact, was mainly a projection of this aggression onto the father. The aggression against the father, which was somewhat special in this case, was constantly provoked by frustration of the need to be loved by the father—a rather distant and forbidding man who made a poor substitute for the mother who had died when the subject was seven.

This, of course, is but a fragment of the whole dynamic lattice—to use Cattell's term. For example, fear of weakness finds other expressions besides self-assertion and has other sources besides passive homosexuality, and similarly for the various other contents. But it is enough to indicate the sort of thing that some of us find most interesting and particularly important. My interest now is in trying to spell out what it is that distinguishes this kind of dynamic determination from others that are common in personality theory.

It follows from what has been stated above that the structural concepts of *central* and *inner* are not helpful, since they may be—and are, with justice—applied in cases where the determination is of quite different kinds. We cannot say that we are talking about things that are more central or more inner. Things which lie at the bottom of our vertical lattice, or farther back in our chain of determination, may relate di-

rectly to processes in the outer layer of the personality, and have their direct representatives in behavior.

We can, of course, speaking now in terms of dynamics and not in terms of structure, call such things as passive homosexuality or castration anxiety "deep" or even deeper needs, but we will not convey very much to various other psychologists who use precisely these words to describe quite different motivational factors.

Actually, from the dynamic point of view, there is no good reason why we should draw the dynamic lattice vertically, with underlying and overlying contents. We could as well draw it horizontally, with the more genotypical things on the left and the more phenotypical ones on the right.

It is well to remind ourselves that self-assertive behavior might have quite different determination than that illustrated by our case. For example, it might have come into being as a transformation, in accordance with reality, of infantile omnipotence fantasies that have blossomed under a long history of gratifications, and reached full fruition as a means for promoting various important goals of the person. It would not be difficult to show, in some instances, that a conscious or preconscious tendency to self-assertion was the motivating force behind diverse specific performances carried out over long periods of time; and in these cases many psychologists would not hesitate to call the motive deep.

What I am leading up to is probably obvious by now: *unconscious* is the crucial concept for distinguishing the kind of dynamics illustrated by the case mentioned above. I mean, of course, unconscious in the psychoanalytic sense; a process is unconscious that is prevented from becoming conscious by the other inner forces.

For the study of unconscious proc-



esses we do not have to rely upon subjective reports. (It should be added at once, however, that since distortion of the symbolic process is of the essence of unconsciousness, it seems slightly ridiculous to ignore what people say.)

For those who do not like too much talk, but prefer to take their cues from the more overt kinds of behavior, the way is open to the objective study of unconscious functioning. Dollard and Miller (6) have recently performed an important service by translating Freud's description of unconsciousness into the familiar terms of behavior theory.

Behavior that is motivated by unconscious forces is characteristically inflexible and stereotyped, automatic and obligatory, repetitive in disregard of the situation or the consequences of the behavior. Unconscious motives are marked primarily by insatiability and resistance to modification through experience; they tend to be unresponsive to pleasure or pain, rewards or punishments, logic or argument. As indicated earlier, unconsciousness is a failure to make the connections, usually among symbolic processes, that adaptation would normally require.

Resistance to change, though surely a feature of unconscious processes, is not by itself a dependable criterion. Other processes—such as well thought-out convictions, patterns of role taking in highly structured social situations, even certain overlearned habits—may be equally resistant to change. The crucial question concerns the conditions of change.

Unconscious motives, usually, yield only to special devices for making them conscious; other enduring processes may be modified by new experience including new argument, changes in the social situation, heroic de-conditioning—things which may be just as difficult to arrange as an expansion of the boundaries of consciousness.

Unconscious motives, characteristic-

ally, are laid down relatively early in the individual life, but this does not distinguish them from other motives. We may find in the individual adaptive, controlled, and flexible motives which in their essence have been there more or less from the start. In unconscious motives the infantile elements may be said to have suffered early repression and to have been brought over, unmodified, into the present. Of other motives it may be said that, just as early, they became controlled or adapted to reality, and that they have persisted fundamentally unchanged just because of their flexibility.

In sum, then—and now I would refer back to the above list of meanings of surface and depth—in sum, with respect to the kind of motives that I have called most interesting and particularly important, one might say that they tend to be deep, inner, hidden, basic, central, integrating, genotypic, resistant to change, and originating early in the life of the individual. But none of these characteristics apply only to this class of motive. One may even argue for their universality in the species or for their relatively heavy admixture of biological elements without approaching very close to their unique character. What does distinguish them is their unconsciousness, and the special dynamics that proceed from this state of affairs.

Some of my colleagues may conclude that I have ceded too much ground to nonpsychoanalytic or even nondynamic psychology. My answer is that it is best to face facts and that, happily, we can afford to. Although functional autonomy, or something like it, is no doubt here to stay, we do not need to yield an inch on the line that unconscious motives, with all their elaborate transformations and counter cathexes, are to be found universally in the human personality.

I have elsewhere (16) termed these

processes "non-extinguishing motives," in an effort to take the discussion of them out of the context of "the neurotic" versus the normal. We might go so far as to admit, however, with Kubie (10), that we are dealing here with "neurotic potential"; something which is surely the stuff of which neurosis is made, given the critical external conditions, but which may also be the major motivating force in creative and socially constructive activity. I believe we showed in *The Authoritarian Personality* (1) that this potential is, in the culture of today, a determinant of social attitudes that are very widespread.

In research on personality the object of investigation is the psychological organization of the individual. We distinguish parts or subsystems and try to understand them in their relations to the whole system. We rarely succeed in getting a purchase on *all* the parts, but we may certainly aspire to a grip on *all the kinds* of parts. Among the kinds of parts or subsystems are unconscious ones, and no study of personality that claims to be thoroughgoing can leave them out of account.

The detection and estimation of unconscious motives and mechanisms proceeds according to the same general logic that holds for the investigation of any other subsystems of personality. We observe consistent trends in behavior, and on this basis conceive of regularities of functioning in the personality; we then design experiments to test whether these hypothetical regularities of function express themselves in predictable ways in specified situations.

In spelling out the behavioral indices of personality systems, we make progress both by clarifying our conceptions of these systems, as this paper has attempted to do, and by sharpening our observations. Faced with the fact that similar behaviors may have different sources, we should ask, "How similar

are those behaviors?" For example, it seems quite possible that we might learn in time to distinguish, by direct observation, self-assertive behavior that is overcompensatory from self-assertive behavior that is conscious and adaptive. We would thus define two self-assertive variables, prepare a manual describing their common manifestations, and have them independently rated—preferably by different raters. And the same could be done for various other unconscious and conscious motives.

In determining the indices of unconscious factors we shall have more difficulty with our criterion measures than with our tests or objective instruments. Our Minnesota colleagues have long since delivered the challenge that any personality variable, of any kind, for which criteria can be established can be measured by a scale such as the MMPI. Paul Meehl (12) has recently made it clear that the challenge holds for castration anxiety, repressed sadism, unconscious Oedipus wishes, and what have you. I should like to see this challenge accepted. I should like to see an enormous amount of effort given to the determination, in more or less normal individuals, by intensive clinical studies, of a wide variety of hypothesized unconscious motivational tendencies. Then, if the empirical test developer wants to use these as criterion measures, I should be quite happy. And I would predict considerable success for him. One might very well doubt that he would do any better than someone who used indices from projective techniques or selected items according to theory—but let us wait and see.

Anyway, it would certainly be a fine thing for personality research if some of our vast resources for developing empirical tests could be devoted to studies in which the criterion measures are inner motivational forces rather than the usual external performances.

If this were done, it should be possible in time to obtain measures of enough efficiency so that it would be feasible to study numerous variables of personality at the same time, and each in relation to the others. It is mainly because the clinical case study has been able to do precisely this that we owe most of our fundamental knowledge of personality organization to this particular technique of investigation. Today, when we appropriately accent objectivity and the testing of hypotheses, we cannot be content with clinical material alone, and we know that it is not necessary for the individual whom we study intensively to be a patient in psychotherapy. What is called for is personality assessment of the kind introduced by Murray (14) at the Harvard Psychological Clinic in the early 1930s. Personality assessment, in the beginning, was pure research directed to the personality as a whole.

It seems that research of this kind has to be regarded as a casualty of World War II. Personality assessment was adapted, in the OSS, for the attempted prediction of future performances, socially defined and socially important. This has been the major pattern of large-scale researches into personality since that time. That personality assessment has not been highly successful in its efforts to predict future performances should not have surprised us, since any socially important behavior depends on a great many factors besides those of personality. But the unfortunate thing, from the point of view of one interested in developing a science of personology, is not so much the embarrassment of being overextended as it is the general shift of emphasis from the study of personality to the study of behavior—usually practically important, socially defined behavior. The study of social behavior per se—the study of such be-

havior as prejudice, staying in college, performing well in professional training—with a view to discovering all its determinants is social psychology. Social psychology is a fine discipline; the study of behavior of the kind just indicated is interesting, important, significant; but it is not personology. Certainly personality factors enter into the determination of social behavior, but the student of personality should not be troubled by the efforts of the social psychologist to win more ground for his own discipline. If, for example, it should turn out that in a given group of subjects prejudice was more a matter of the social situation and less a matter of unconscious defense mechanisms than the personologist had thought, he does not redouble his efforts to predict prejudice but he asks, instead, "How, then, do these subjects manage their unconscious impulses?"

Social psychology, and general psychology, have been enriched by personology. But personology has made this contribution at considerable expense to itself. It can be replenished only by thought and investigation directed to personality for its own sake.

#### REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, ELSE, LEVINSON, D. J., & SANFORD, R. N. *The authoritarian personality*. New York: Harper, 1950.
2. ALLPORT, G. W. *Personality: a psychological interpretation*. New York: Henry Holt, 1937.
3. ARON, BETTY. *A manual for analysis of the thematic apperception test*. Berkeley, Calif.: Willis E. Berg, 1949.
4. BERTALANFFY, L. VON. Theoretical models in biology and psychology. *J. Pers.*, 1951, 20, 24-38.
5. CATTELL, R. B. *Personality*. New York: McGraw-Hill, 1950.
6. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.

7. FREEDMAN, M., LEARY, T., OSSORIO, A., & COFFEY, H. The interpersonal dimension of personality. *J. Pers.*, 1951, **20**, 143-161.
8. FRENKEL-BRUNSWIK, ELSE. Psychoanalysis and personality research. *J. abnorm. soc. Psychol.*, 1940, **35**, 176-197.
9. FRENKEL-BRUNSWIK, ELSE. Motivation and behavior. *Genet. psychol. Monogr.*, 1942, **26**, 121-265.
10. KUBIE, L. S. The fundamental nature of the distinction between normality and neurosis. *Psychoanal. Quart.*, 1954, **23**, 167-204.
11. LEWIN, K. *Field theory in social science*. New York: Harper, 1951.
12. MEEHL, P. Wanted: a good cook-book. Presidential address, Midwestern Psychol. Ass., April 29, 1955. (Dittoed)
13. MORGAN, C. D., & MURRAY, H. A. A method for investigating phantasies: the thematic apperception test. *Arch. Neurol. Psychiat.*, 1935, **34**, 289-306.
14. MURRAY, H. A., et al. *Explorations in personality*. New York: Oxford Univer. Press, 1938.
15. SANFORD, R. N., et al. Physique, personality and scholarship. *Monogr. Soc. Res. Child Developm.*, 1943, **8**, No. 1.
16. SANFORD, R. N. Family impact on personality: the point of view of a psychoanalyst. In J. E. Hulett & R. Stagner (Eds.), *Problems in social psychology*. Chicago: Univer. of Illinois Press, 1954.
17. TOMKINS, S. S. *The thematic apperception test*. New York: Grune & Stratton, 1947.

(Received November 11, 1955)

## CONSCIOUSNESS AS A REGULATORY FIELD: A THEORY OF PSYCHOPATHOLOGY

REX M. COLLIER

*The Veterans Administration*<sup>1</sup>

In a previous paper which provided an initial statement of the writer's regulatory theory of consciousness, reference was made in brief detail to the physiological gradient (3, p. 273), and it was suggested that consciousness, generated at the organic levels of highest metabolic rate, becomes regulatory in psychological processes. This initial statement applied only to nonstress processes. The purpose of the present paper is to develop the theory in the direction of stress-defense dynamics, and to make applications to selected basic issues in psychopathology.

Consciousness is considered here as a field of interacting energies or processes whose chief function is to increase the flexibility of adjustment. The interaction of these processes, together with their resultant outcomes, seems appropriately conceptualized under the general, biologically derived term "regulation." While no attempt is being made to present a complete theory of consciousness, it is proposed that consciousness emerges as a basic property of the highest level of metabolic activity in the physiological gradient field. A second, related proposition is that consciousness provides the field conditions on a psychological level for adaptive regulation of behavior. The proposition that regulation is the basic function of consciousness invites one immediately to follow out certain derivable implications. This is a formidable task and one cannot hope by initial statements to answer all questions nor satisfy all readers. To

use the term consciousness is to call up by association a morass of age-old problems and controversies. The writer proposes to bypass most of these by remaining primarily concerned with the implications of consciousness as a regulatory field.

Adaptive regulation is characterized, principally, by flexibility of responding to changes in stimulus patterns either from within or without the organism or from a combination of both. Thus, the more any process is admissible or accessible to the conscious field, the greater are its chances of entering into mutually modifying interaction with other relevant processes that are also available to consciousness. The behavior resultants, likewise, have increased chances of thereby being more representative of a total pattern of needs than an impulsive response would be. Conversely, processes occurring outside the conscious field or at less clear levels are more rigid, and their degree of inflexibility is assumed to vary with the degree to which they and their associated processes are inadmissible to the clearer or broader levels of the field. Rigidity, compulsiveness, and lack of flexibility indicate a limited conscious field in relation to such responses and a large amount of unconscious direction.<sup>2</sup> The variations, limitations, and pathologies of the postulated regulatory field of consciousness become, collectively, a set of problems of basic importance in psychology. It is suggested further that uniting the concepts of regulation and

<sup>1</sup> From the Jefferson Barracks Veterans Administration Hospital, Missouri.

<sup>2</sup> The reader is invited to compare G. W. Allport's concept of relative freedom with this concept of flexibility (1, pp. 84-85).



consciousness may provide for both a greater integration between psychology and biology and for a more systematic approach to the problems of personality.

Viewed in the manner indicated, consciousness becomes a progressive achievement of utmost importance in the evolutionary process. Its role is no longer that of a curious epiphenomenon, but rather a dynamic set of conditions enabling increased flexibility and appropriateness of adjustment. To recognize this functional role of consciousness is to restore to psychology, in new form, a set of problems and approaches that we have too long denied to ourselves.

This assumption that the conscious field is regulatory in the adjustive sense implies also the assumption that the efficacy of regulation is dependent upon the current magnitude of the field. By magnitude<sup>3</sup> it is implied that the conscious field may vary in the degree to which processes relevant to current issues may be accessible to the field. It is assumed that the more permeable the field boundaries are to relevant processes the broader and the more inclusive is the regulatory frame of reference. Blocking, resistance, repression, etc. represent decreases in permeability of field boundaries and essentially constrict the degree to which processes potentially relevant to an issue are available. As Allport has implied (1, p. 85), a person who is limited on one basis or another to only one solution of a problem has only one degree of freedom. It is assumed further that reductions in magnitude of the conscious field reduce its regulatory potentialities, with coincident reductions in flexibility of approaching problem issues and in the chances that behavioral outcomes are

of long-term benefit to the individual. Thus, it immediately becomes conceivable that regulatory effectiveness may be very greatly reduced, given sufficient stress along with maladaptive types of defense. The regulatory effectiveness of the conscious field is, therefore, limited. The conditions producing variations in the degree of limitation have, in part, just been implied, and they provide the major subject matter for the remainder of the discussion.

#### BASIC DYNAMIC CONCEPTS

In order to achieve brevity and simplification, it is proposed that basic dynamics relevant to the regulatory conscious field may be defined in terms of the following concepts: (a) stress, (b) defense system, (c) stress-tolerance threshold, (d) ego strength, and (e) boundary permeability of the conscious field. These are defined as follows:

*Stress* is a broad term including a wide range of psychological, physiological, and physical factors, all of which have the common element of placing demands on the organism's capacity to react to a degree that usually becomes uncomfortable or threatening.

*Defense system* is an interrelated set of psychological reactions, protective in nature, which either increases the psychological distance of the threatening issue (primary segment of defense system), or disguises or distorts the threat so that in its new form it becomes more acceptable (secondary segment of defense system). By the concept of system it is hoped that we may advance from the outmoded atomistic concept of single "mechanisms" to a more molar and internally consistent concept.

*Stress-tolerance threshold* is described in terms of the amounts of stress necessary to activate to a defined degree the individual's defense system. Figure 1 will aid in identifying three such thresholds.

<sup>3</sup> Magnitude as it is used here implies a kind of spatial model. While the model need be neither circular nor spherical, it could, for ease of analogy, be either.

*Ego strength* is the ability of the individual to accept into clear consciousness potentially stressful or disturbing material and to assimilate it constructively. While ego strength has traditionally included a number of associated potentialities such as orientation, reality testing, and ability to integrate, the writer proposes that the above definition is basic to all of these.

*Boundary permeability of the conscious field* refers to the degree to which the conscious field is accessible to segments of experience, past or current, which are related to processes operating in the field at any given time.

Since these conceptualized variables are dynamically interrelated, variations in one have resultant effects in the others. It seems difficult, if not impossible, to deal with one in isolation from the others, and to define any one leads to some degree of circularity.<sup>4</sup> Taken together they lead to the *central concept of stress-defense dynamics*. For some purposes one might define stress objectively or operationally in terms of volts, decibels, foot candles, or units of pressure. The effects, however, would vary widely because of varying thresholds of tolerance. Situations that some individuals accept and manage with ease would cause others varying degrees of discomfort. It is necessary in the present approach to deal conceptually, at least, with the experienced aspects of stress. When experienced stress such as discomfort, tension, and anxiety exceed the tolerance threshold, the individual reacts defensively according to his habitual system of defense. Thus the concept of stress becomes immediately related to tolerance and defense. Ego strength defined as above is a central factor in determining how stress is perceived. "Good ego strength" goes along not only with less need for de-

fense but with higher stress-tolerance levels. What constitutes stress in its relative effect is directly related to degree of ego strength. Since ego strength has been described as ability to accept into awareness potentially stressful material and to assimilate it constructively, tolerance threshold and needs for defense become dynamically related not only to this concept but to each other. Two segments of the defense system were implied above. These must be distinguished in more detail.

#### PRIMARY DEFENSE

While all defenses have at least one end in common, namely, to reduce psychological discomfort or anxiety and to restore some degree of "peace of mind," the most primitive, elemental, and perhaps the most frequently utilized defense at all levels of animal behavior is flight. The internalized equivalent of this elemental, primitive, and frequently used defense is a constriction of the field of consciousness. While flight may be expressed behaviorally through transportation activity from one location to another, it can be expressed on an experience level by reducing the psychological nearness of the threat. Thus, if the individual cannot successfully rid himself of the threat by fleeing, he may attempt reduction of his anxiety by avoiding thoughts of it. Some relief is gained by variations of this self-imposed constriction of the conscious field through selectively shutting out tension-producing thoughts, feelings, memories, and meanings. This segment of the defense system the writer has chosen to call primary because of its elemental character. Many gradations of this constricting process or primary defense can be noted and illustrated. For example, one who whistles in the dark or sings to avoid thinking, or feeling threatened, represents a relatively mild example. At the other extreme, deep, and extensive operations of denial and

<sup>4</sup> Cf. A. H. Maslow's treatment of cause-effect circularity in his discussion of the holistic dynamic point of view (6, p. 27 f.).

repression relegate large segments of experience beyond relatively impermeable barriers from the conscious field. The tendency is illustrated by the following verbalizations often heard in a clinical situation: "This is too horrible to think about." "The best way for me to get over my war experiences is not think about them." "Thoughts like that I keep out of my mind as far as possible." "I have tried my best to forget it."

The naive implication in the use of this kind of defense is that to put a thought or feeling "out of mind" is to annihilate it. Were the assumption entirely valid, the solutions of many problems would be simple. That this is not a simple, effective solution is apparent from repressed impulses which seem to be buried, but which generate a variety of ghosts to haunt a troubled consciousness. Repression is not to be thought of as completely synonymous with the present concept of self-imposed restriction of the conscious field. Repression is found at one extreme of a constriction continuum. The defense of restriction or constriction which is represented as a primary defense may be expressed over a continuum of varying degrees of intensity and extensity.<sup>5</sup>

The assumption that the conscious field is regulatory in the adjustive sense implies also the assumption that the efficacy of regulation is dependent upon the potential magnitude of the field. It will be recalled that magnitude here means the degree to which processes relevant to an issue are accessible to consciousness. When stress rises above the threshold of tolerance, and is either persistent or increases in intensity, the continued utilization of primary defense exacts a price which is paid in terms

of reduced effectiveness of regulation. That is to say, the degree of dominance of the conscious field is dependent upon the degree to which experience relevant to the needs of the moment is available. It follows, therefore, that whenever the primary defense is used, some degree of crippling of the regulatory function of the conscious field results.<sup>6</sup> Furthermore, if the threat persists and the individual is committed to or utilizes only the primary defense, the crippling effect becomes progressive. The end result is an adaptive regulatory field so impoverished, inadequate, or impotent that serious disorganization in thought, feeling, and behavior occurs. Behavior becomes dominated by impulses that are as rigid as their degree of dissociation from consciousness. The impulses are no longer organized or regulated in terms of comprehensive, representative, and realistic need patterns. Rather the impulse of the moment becomes of chief importance, whether generated from the id or from the residuals of experience that form the superego. Obviously typical orientations in regard to self, persons, places, and time may be seriously affected. Hence we have some form of what has been called a psychosis. When, therefore, stress increases sufficiently and the balance of the defense system tends chiefly to the primary segment, the end result is self-annihilation, which may be either suicide or psychosis. That Anna Freud arrived at a somewhat similar conclusion by her observations is indicated by the following: "But repression is not only the most efficacious, it is also the most dangerous, mechanism. The dissociation from the ego entailed by the withdrawal of consciousness from whole

<sup>5</sup> To be consistent in terms of spatial models we should probably indicate a plane rather than a line (continuum) if we assume that both intensity and extensity describe primary defense variations.

<sup>6</sup> It is also implied that narrowing the conscious field by specific kinds of suggestion, e.g., hypnotic suggestion or propaganda, certain forms of advertising appeal reduce the degrees of freedom of the individual even though only temporarily.

tracts of instinctual and affective life may destroy the integrity of the personality for good and all" (5, p. 54).

#### SECONDARY DEFENSE

The defense system is generally recognized as having a degree of stability greater than that implied by the operation of primary defense by itself. Behavioral disturbances often occur, though they seldom end in psychoses; neuroses are recognized as more common than psychoses, and the progressive crippling resulting from a predominant use of primary defense is avoided in most instances of continued stress. The assumption here is that secondary defenses are developed and utilized for stabilizing and arresting the otherwise progressive crippling trend. Thus a homeostatic balance (4, 9) preserves the self structure which continues to function with a degree of impairment. The proposition is offered that the degree of impairment becomes a resultant of at least the following factors: (a) intensity and persistence of stress, (b) the regressive stage at which the secondary defenses were developed and elaborated, (c) the adequacy of the secondary defenses, and (d) the degree of ego strength remaining.

As a first approximation to the distinction between primary and secondary defenses it is proposed that all the typically described ego defenses, excluding repression and denial, may be classed as secondary. Rationalization, compensatory devices, reaction formation, ritualistic behavior, conversion symptoms, etc. are included as secondary defenses. Since it is assumed that, generally, primary defense tends to operate first with the secondary process entering later, secondary defenses appear also to have the character of being restrictive and repressive. This is more apparent than real. Such impres-

sions are the results of incomplete analyses of the individual's defense system. The secondary defenses operate in a dual capacity, first as an aid to the primary defense component in maintaining the status quo of the current constriction of awareness with regard to relevant issues, and second to cushion the threats of additional stress from without and thereby reduce the need "to leave the field." A rationalization, for example, contributes on the one hand to a maintenance of self-limitation or repression already present, and on the other hand to an interpretation of the threat that allows the individual to deal with the situation with some degree of relevance—a degree assumed to be usually unattainable under those conditions without the rationalization. The same would be said of a ritualistic compulsion. The primary defense component having already operated to a degree, the secondary component (the ritualism) not only maintains the status quo in degree of repression, but also reduces the intensity of threat and further need for operations of primary defense. The experimental production of disturbed behavior (neuroses or psychoses?) in animals could test this point in part. Hypothetically, animals which develop a ritualism should show less of the symptoms of regression than animals which do not develop such a secondary defense.

Classifying the traditional ego defenses other than repression and denial as secondary defenses will probably not prove entirely satisfactory. Traditional defense terminology was developed at a time when, as "mechanisms," these defenses were thought of as operating with considerable differentiation and identity. It seems more appropriate to hypothesize a defense system characterized by a dynamic pattern of relationships between primary and secondary components. The primary component tends

progressively to cripple the control field and has attributes of malignancy, while the secondary component tends to arrest and stabilize the pathological process and is more benign. Primary defenses insidiously destroy homeostatic stability while seemingly contributing to a kind of balance. The process of mastering anxiety under conditions where the primary component is dominant is usually something less than successful. The struggle to maintain equilibrium perhaps continues to the last, even though adaptability progressively decreases.

Traditional approaches in understanding behavioral disturbance have emphasized stress factors along with varying degrees of attention to constitutionality. The implication of this theoretical formulation is that the individual is not overwhelmed simply by the intensity of stress along with some vaguely defined constitutional weakness but rather, and perhaps particularly, by his mode of habituated defense. A correlative assumption is that one of the most important features of the individual's "life-style" pattern is his system of defense and the kind of balance between primary and secondary components. Knowledge of the factors determining the individual's predispositions to a particular balance of his defense system would be of great value in personality theory and in clinical applications.

#### PRIMARY DEFENSE AND THE CONCEPT OF HYPEREXIS

After having arrived, on the basis of clinical observation but with some trepidation, at a limited application of homeostatic principles, the writer discovered Richards' "Homeostasis versus hyperexis, or St. George and the dragon" (7). Richards insists, contrary to Cannon (2), that the body is not always wise, rather it can be "stupid, egregiously, calamitously stupid" (7, p.

291). He finds numerous evidences from various kinds of pathology that the body often develops an excess type of response. That is, "in trying to be homeostatic in one direction, the body finds that it has been most un-homeostatic in another" (7, p. 291). For this excess response he offers the term "hyperexis," and argues that this is as much a part of the "scheme of things" as homeostasis.

The hyperexis which Richards observes in tissue pathologies seems also to be expressed in psychological patterns where emphasis on the primary defense component results in a progressive reduction in the flexibility and effectiveness of the consciousness regulatory field. Overcompensation, overgeneralization, and the so-called "neurotic paradox" may be found to be examples of psychological hyperexis in their self-defeating results. Thus what looked at first like a homeostatic response becomes unhomeostatic; what looked also, at first, like wisdom of the unconscious becomes "calamitously stupid," since through the psychological hyperexis there results in some degree—sometimes final—self-annihilation. It is rather the secondary defense component that operates to restore and preserve homeostasis. This primary-secondary defense dynamic shows why it is so dangerous to rob an individual of his secondary defenses before he is ready and able to relinquish them. This issue should be clearer following the discussion of Fig. 1.

#### NEUROSIS AND PSYCHOSIS

Controversy currently characterizes much of the diagnostic thinking with regard to psychoses versus neuroses. Some argue that neurotics never become psychotic and, therefore, the two disorders remain distinct. Others assume a kind of continuum of psychopathological severity, with the psycho-



ses at the high end of the severity continuum. Rosen has made the following comment on this issue: "We regard the neurotic and psychotic as different only in degree of stage of illness; one is simply sicker than the other. I do not agree with those colleagues who maintain that because an individual becomes neurotic he is therefore unable to be psychotic. Insofar as we all have an unconscious we are all potentially psychotic" (8, p. 6). Rosen has also pointed out more specifically a kind of continuum with regard to psychotic and neurotic variations.

In writing the papers contained in this volume it became increasingly difficult for me to use the term schizophrenia when what I had in mind was psychosis. When I saw enough patients who had been diagnosed "manic depressive psychosis" and were not schizophrenic, I began to realize that psychotics do not differ in kind but only in degree. I became more sure of my ground when in the course of recovery from schizophrenia they again became manic depressive and went on from there to neurosis. I cannot reconcile the different kinds of psychosis with Freud's concept of the unconscious (8, pp. 178-179).

Neither a complete distinction between neurosis and psychosis nor a simple continuum as suggested by Rosen seems appropriate. Psychopathology can be relatively mild but have a psychotic-like character. When this occurs, the clinician may recognize the case as different from a neurotic type of reaction but indicate that even though the symptoms appear mild they could easily become serious. The person is recognized as in good contact, but it will be felt that if current stress continues or increases, the prognosis is poor and a psychotic reaction could easily be induced. On the other hand, the clinician often sees patients who may be much more incapacitated than the type just referred to, but who are described as neurotic. The prognosis here usually does not anticipate psychosis. Some clinicians, however, insist they have seen

such cases become psychotic when the progress of the neurotic pathology could not be arrested. Insofar as these are valid observations, they fit the dynamic schism presented below in Fig. 1.

In terms of this theory a psychosis is an end result of continued or habitual use of primary defense under stress; it may also be the result of breakdown of the secondary defenses which then revert the individual to the progressive crippling or malignancy of the primary defense process. A neurosis, then, becomes that resultant symptom pattern which is elaborated by the development of secondary defenses. The writer does not propose by the use of the terms "psychoses" and "neuroses" to seek validation for their common use. Considering the current dissatisfaction with the terms, better ones may be developed. The purpose is rather to relate what is commonly recognized under these heads as dynamically related to the manner and degree in which the self regulatory system is crippled or stabilized by a pattern of defense characteristically used by the individual under stress. If this formulation is correct, then in diagnosis more attention should be given to eliciting and understanding the individual's defense system. A knowledge of why some individuals remain dependent on primary defenses while others elaborate secondary defenses, as needed, should add to our understanding of whatever distinction there may be between the psychotic trend and the neurotic patterns.

Figure 1 represents graphically an approximation of the interactions of several of the major dynamic factors theoretically operative in psychopathology. The  $x$  axis represents increasing theoretical increments of stress. The  $y$  axis represents the two components of the defense system. The secondary defense component increases in ascending from the  $x$ ,  $y$  intersection or origin, while the primary defense component

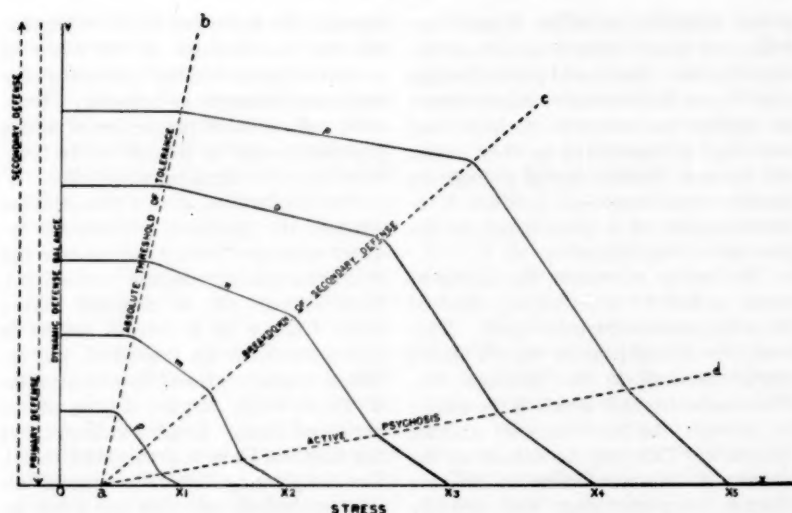


FIG. 1. Graphic representation of the theoretical stress-defense functions (see text for explanation).

increases on the same axis in the descending direction. Each point on the y axis, then, represents a balance or a kind of algebraic summation between the primary and secondary defense factors. The line *ab* is labeled "absolute threshold of tolerance," and is defined as the minimum degree of stress necessary to activate the individual's defenses. To the left of *ab*, increments in the stress factor are theoretically insufficient to activate the defense system. The slope of the line *ab* needs a limited interpretation in terms of this two-dimensional diagram. Taken as it stands, it would seem that the more predisposed an individual is toward utilizing secondary defenses the higher is his absolute threshold of tolerance. While this is partly true, the more complete relationship would appear had we included, as a colleague<sup>7</sup> has suggested, a third dimension labeled "ego strength." The concept of ego strength has already been defined as the degree

to which an individual can maintain permeable barriers to the conscious field and admit to awareness potentially stressful material and assimilate it constructively. Were the third dimension, as defined, included, it would be clearer that increases in ego strength reduce the need for any of the defenses. However, given both high ego strength, initially, and stress that persistently exceeds tolerance thresholds, the individual will be inclined not only to use whatever degree of primary defense he must but will be more resourceful in the utilization and elaboration of secondary defenses.

The concept of control in ego strength is included by implication and given a basis for its operation. If the conscious field is regulatory, then permeability of its boundaries to processes relevant to a given issue is necessary for the appropriate distribution of the regulatory effects. Success in psychotherapy is indicated by increased ability of the individual to admit to consciousness re-

<sup>7</sup> Dr. C. Scott Moss.

jected material, to utilize it meaningfully, and to act upon it in a constructive direction. Successful psychotherapy also means the correlative improvement in ability to recognize feelings and meanings of experience as they occur. All therapy, finally, should attempt to provide experiences that facilitate a re-establishment of a more adequate degree of self-regulation.\*

The line *ac* represents the degree of stress at which the stabilizing effect of secondary defense begins to break. This could be thought of as the "Plimsoll mark" referred to by Freeman (4). This point beyond which it is unsafe to increase the load becomes another threshold. This may be defined as the degree of stress sufficient to activate defense symptomatology that severely interferes with social and economic adjustments. With successive increments of stress the primary defense component becomes more active. Or, we can say that when secondary defenses finally become inadequate, the individual quickly reverts to primary defense reactions with only relatively small additional increases in stress. From this point on, further increments of stress are relatively more effective in bringing the individual to a complete break or a psychotic reaction. The line *ad* represents, therefore, another threshold function. This is defined as the amount of stress required to initiate an active psychosis. As with the other threshold functions, this line also passes through a series of inflection points. The stress required to exceed this threshold depends also upon the dynamic balance in the individual's defense system.

The points  $X_{T-1}$  represent terminal thresholds of tolerance. Increases in stress beyond this threshold give no changes in reaction. The increase in stress may be meaningless. Thus, for

example, an individual in active psychosis may be informed of the death of a near relative without producing any noticeable changes in behavior. However, only a small proportion of active psychotics may be thought of as functioning at the terminal threshold.

The graphs *l*, *m*, *n*, *o*, *p* represent essentially the functional relationships between stress and various balances in the primary-secondary defense components. Since it may also be assumed that a given balance in a defense system is characteristic of an individual, the series of graphs represent hypothetical individuals under varying degrees of experienced stress. Graph *l* indicates that the individual who is predisposed to utilize the primary defense component almost exclusively not only has a low series of thresholds to stress, but loses very rapidly any slight tendency he has toward secondary defense elaboration. He reaches the region of active psychosis very quickly (not necessarily in time but in degree of increase in stress) and, since little of the secondary elaboration is observed, he is thought of as being distinctly out of the neurotic group. Graphs *o* and *p* represent individuals who appear to the diagnostician as having one or a combination of neurotic symptom patterns as their behavior approaches the *a*, *c* region. The *n* graph represents individuals whose symptomatology in the *a*, *c* region confuses the diagnostician as to whether he is dealing with a psychotic or a neurotic trend. If the administrative importance of the diagnostic label is subordinated to the value of understanding the individual, the *n* graph ceases to represent an area of confusion but rather an intermediate degree of severity. The more the basic dynamic pattern of an individual is revealed, the less need there appears for the traditional diagnostic classification. Were instruments and methods available for evaluating the defense system and ego

\* This therapeutic purpose is offered in contrast to certain traditionally stated goals of "adjusting the individual to his environment."

strength as defined, we should be able to improve our predictions of behavior in a variety of important situations.

The slopes of the graphs decrease from  $l$  to  $p$  between lines  $ab$  and  $ac$ . This appears to represent the theoretical position that the more an individual utilizes secondary defense the more stress he can absorb before the danger point of excessive overloading is reached. Even beyond this point, the graph implies that he has greater reserves. This is valid, in terms of the theory, providing one does not omit in his thinking the factor of ego strength. This dimension, not included in the graph, is associated with reduced need for defenses in general, but where stress increases sufficiently, ego strength, as defined, would predispose the individual more toward secondary defense elaboration than toward primary defense. The reserves referred to are as much a result of ego strength as they are the type of defense.

#### SUMMARY

The theory that consciousness is dynamic rather than passive and that it is regulatory within limits has been extended and applied to selected aspects of personality. These extensions of theory have included specifically the relationship of defense systems to behavioral deviations. The proposition was made that consciousness be considered a biopsychological field evolved to provide the essential conditions for increasing the complexity and flexibility of adjustment. Consciousness was represented as an interactive field wherein processes (e.g., thoughts, feelings, memories, needs, meanings, etc.) have increased chances for mutual modification. The regulatory effectiveness of this field depends upon a kind of magnitude which is defined as the degree to which all processes relevant to the issue are available.

Defense reactions were described in terms of a dynamic defense system, the two chief components of which are primary and secondary defenses. Primary defenses are the more elemental and tend to be a psychological flight from stress through a constriction of the conscious field. Continued use of this defense impairs the regulatory effectiveness of the conscious field. Secondary defenses, incompletely defined as the typical ego defenses other than repression and denial tend to stabilize the otherwise progressive impairment under stress. Psychoses result from excessive use of primary defense under continued stress, while the neuroses become the elaboration of secondary defense.

The concepts of stress, defense system, stress-tolerance threshold, ego strength, and boundary permeability are defined in ways to provide systematic relatedness with this approach. A suggested definition of psychotherapy is included but will be elaborated later.

#### REFERENCES

1. ALLPORT, G. W. *Becoming*. New Haven: Yale Univer. Press, 1955.
2. CANNON, W. B. *Bodily changes in hunger, pain, fear and rage*. (2nd Ed.) New York: Appleton-Century-Crofts, 1929.
3. COLLIER, R. M. Outline of a theory of consciousness as a regulatory field: preliminary statement. *J. Psychol.*, 1955, 40, 269-274.
4. FREEMAN, G. L. *The energetics of human behavior*. Ithaca, N. Y.: Cornell Univer. Press, 1948.
5. FREUD, ANNA. *Ego and mechanisms of defence*. New York: International Universities Press, 1946.
6. MASLOW, A. H. *Motivation and personality*. New York: Harper, 1954.
7. RICHARDS, D. W. Homeostasis versus hyperexis: or St. George and the dragon. *Sci. Mon. N. Y.*, 1953, 77, 289-294.
8. ROSEN, J. N. *Direct analysis*. New York: Grune & Stratton, 1953.
9. STAGNER, R. Homeostasis as a unifying concept in personality theory. *Psychol. Rev.*, 1951, 58, 5-17.

(Received May 17, 1955)

## FRUSTRATION THEORY: RESTATEMENT AND EXTENSION<sup>1</sup>

NORMAN R. F. MAIER

*University of Michigan*

In presenting a new point of view one must of necessity differ with the thinking of others. The perception of these differences may also vary considerably, so that a communication problem is introduced. The first step in resolving the disagreements resulting from an initial presentation of a viewpoint seems to be to understand the sources of the disagreements, assuming that emotional factors can be regarded as beside the point of a scientific discussion. Some of these differences are purely semantic; others are due to the emphasis the writer puts on the aspect of his views that are new, so that readers may assume that he has neglected or overlooked other important relationships; while the remainder are due to alternate ways of interpreting the same phenomena. The first two sources of disagreement can conceivably be settled by further writing and discussion, but the third needs to be resolved by research. Before new research can adequately settle the relative adequacies of differing points of view, however, the basic differences must be separated from differences resulting from faulty communication.

For the sake of progress in science it is also desirable to consider a fourth source of difference. No point of view remains static, and it is questionable whether an author can reinstate the meaning he originally intended when he later elaborates his position. His views must grow as a result of additional years of speculation, and pointed criticisms

help him clarify his distinctions and stimulate him to come to deal with relationships that he may have neglected. Failure to communicate these later developments may unwittingly lead to misunderstandings and disagreements.

The frustration theory of the writer (16) seems to be a case in point. Since its presentation in book form, there have been many descriptions, verbal statements, and critical research reports that make assumptions about the theory that seem unjustified to the author of the theory. Without attempting to determine which of the causes mentioned above makes for the discrepancy between the thinking of the author and that contained in restatements of the theory, it seems that some clarification and sharpening of distinctions is in order. It is hoped that such an analysis will clarify the thinking of proponents and opponents, obviate controversy due to faulty communication, and point up crucial experiments that can be formulated. Experimental findings will be used either to separate alternate points of view or to indicate the points at which alternate views are tenable. Since the theory of frustration cuts across learning, motivation, therapeutic processes, and personality theory, the concepts involved need to find a proper place in a general behavior theory.

### CLARIFYING QUESTIONS

#### *Is a Frustrated Person Without a Goal?*

The writer's book title, *Frustration: The Study of Behavior Without a Goal*, may be taken to mean that a frustrated person has no goal. This was not the intent in selecting the title. The

<sup>1</sup> I wish to acknowledge my indebtedness to my colleagues Professors R. W. Heyns and E. B. McNeil, whose critical comments and constructive suggestions have greatly improved this paper.



thought that the title was meant to convey was that a frustrated person's *behavior* is without a goal, i.e., the behavior sample under discussion lacks goal orientation. Suppose someone steps ahead of another who is in line waiting to purchase a railroad ticket. An argument ensues, followed by a fight. Both men are arrested and neither makes his train. However, the need for a railroad ticket remains for each, and the more intense the need is, the more frustrating is the interruption. The fighting, however, is not ticket-getting behavior. According to the theory, the fighting is frustration-instigated behavior. This position differs conceptually from a viewpoint that states that a new goal replaces the one of obtaining a ticket. Examples of possible substitute goals in this example are (a) obtaining relief from tension, (b) demonstrating superiority, (c) teaching a bully his place, and (d) doing injury to an enemy. Even though relief or satisfaction may sometimes result from such aggression, this fact should not be accepted as proof that the desire either for relief or for these satisfactions was an instigator. If such causation exists, it must be shown that it is an essential requisite to such behavior. According to frustration theory, adaptive or need-satisfying behavior is not assumed. Motivation therefore is separated from causation as an explanatory concept. Research is needed to support or reject these basically differing assumptions.

#### *Is the Frustration-Fixation Relationship Circular?*

Circular thinking in this instance would seem to run as follows: frustration produces fixation, and the appearance of a fixation in an animal indicates that it has been frustrated. To the extent that this relationship has been communicated, the author has been lax in that he has failed to clarify the process

by which he reached certain conclusions. An attempt to reproduce the logical steps which lead to the qualitative distinction between behaviors produced by frustration and behaviors produced by motivation learning is given below.

1. When rats were placed in an insoluble discrimination problem situation, all of them adopted a specific response of some sort (usually a position response), and they were persistent in executing this response when forced to make a choice. The consistency with which animals executed the response selected was a new experience to the observers, but this phenomenon did not suggest fixation. What was important was the fact that when a soluble problem subsequently was presented to the animals, some abandoned their position responses and adopted the discrimination response within standard learning limits (less than 100 trials), whereas the majority never learned (600 trials) and instead persisted, without a single variation, in their initial responses. The bimodal character of these learning scores was striking, particularly the fact that the majority of animals could not learn the problem at all.

2. In order to determine whether the failure of these animals to learn was due to the basic strength of the first response, control groups were added. These groups were required to learn an initial response that matched the responses spontaneously adopted by the experimental groups in the insoluble problem situation. When later tested on the discrimination problem, it was found that the majority of them learned the new problem in standard time, but a minority failed entirely. (It was the practice of earlier experimenters to discard the few animals that failed to learn problems of this kind.)

The following facts needed to be explained: (a) unexpected and apparently unreinforced responses of unusual per-

sistence occur spontaneously; (b) the number of animals having such responses varies with the situation, there being few such animals in simple learning problems and many in insoluble problems; and (c) animals that learn and those that fail to learn form two discontinuous populations.

The following hypotheses were developed to explain the findings.

1. Responses that persist when adaptive alternatives are clearly available are different in kind from habits, since habits are subject to change. The new type of response was given the name *abnormal fixation*, so as not to confuse it with the term *fixation* in habit formation. (For experimental purposes the criterion adopted for abnormal fixation was a specific unadaptive response that persisted beyond 200 trials, or more than twice the number required for substituting one habit for another. Tests had shown that animals that fail to learn in 200 trials also fail with more trials.)

2. Frustration causes abnormal fixations. (This hypothesis was suggested by the fact that an insoluble problem should be frustrating to more animals than a learning problem.)

3. Frustration is a physiological mechanism that determines the nature of behavior when the frustration threshold has been passed over. (The possibility was implied that the autonomic nervous system dominated behavior.)

4. Individuals differ in their frustration thresholds.

Thus the concept of abnormal fixation was developed to account for certain experimental findings, particularly the bimodal nature of certain distributions in scores. A comparable term is the concept of a habit. Reinforcement produces habits, and the appearance of a habit in an animal indicates that it has been rewarded after making the response. However, one does not regard this relationship as circular.

Once, however, the casual relationship between frustration and fixation has been established, one can use differences in the number of animals showing fixated behavior as a measure of the number of frustrated rats. This was done in subsequent studies, in which the concept of the frustration-fixation relationship was used as a working hypothesis.

For example, in later studies these conclusions were drawn.

- a. Increases in the length of the insoluble problem period increase the number of animals whose frustration thresholds will be exceeded (24).

- b. An increase in the difficulty of a learning problem increases the number of animals that will be frustrated (21).

- c. Frustration produces fixated behavior, which has a degree of rigidity (simulating compulsive behavior) unlike that found in habits (9, 15, 16, 26, 27, 28).

- d. The presence of bimodal distributions in stressful situations should be examined to determine whether or not frustration has occurred (9, 21, 27, 28). (Bimodal distributions may have a number of causes and frustration is one of them. The appearance of bimodal distributions requires an explanation in any case, and failing to test for it amounts to making an unproved assumption.)

Each of these conclusions is based upon the number of rats showing abnormal fixations, but one can argue that reasoning has been circular only if one questions the original conclusion, which was that the fixations in the initial experimental condition (25) were produced by frustration and not by learning. To question this conclusion, one must go to the facts upon which it rests, not to subsequent studies. The purpose of the follow-up studies was one of extending our knowledge of the frustration mechanism, rather than one of repeating or re-establishing the case for the quali-

tative distinction between the behavior of frustrated and nonfrustrated animals.

### *Is Escape from Punishment a Goal?*

Can we assume that fixations are caused by an animal's attempt to escape punishment on the jumping stand? This represents an alternative explanation of why an animal may persist in an apparently unadaptive response and fails to learn the discrimination problem. According to this view, the animal is rewarded by its response because it escapes punishment, and hence it fails to learn the better way to escape that is offered. What must be explained, however, is the bimodal distribution of scores and why this kind of response differs from reinforced responses. There are many explanations for persistent behavior, and Dollard and Miller (3) make this their task when they reduce abnormal fixations to escape learning. Researches on avoidance learning present evidence in support of the persistence of behavior produced by punishment, and the question may be raised as to whether the frustration-fixation hypothesis is an alternative concept for interpreting the data in their experiments. It is also possible that the interpretations may be supplementary rather than in conflict, particularly if one entertains the notion of two behavior mechanisms: motivation and frustration. Thus punishment may serve either as a negative incentive or as a frustrating agent, depending on its intensity.

### *Is Anxiety Reduction a Goal?*

Anxiety reduction has been previously discussed (19) as an alternative explanation for abnormal fixation phenomena, and it, too, assumes the task of explaining persistent behavior. Anxiety reduction is described as a kind of reward for the fixated response, which means that the response is actually reinforced by

the tension reduction that follows the jump. All that is required is the postulation that animals have a need to reduce tensions. Why an even more rewarding response (one that includes tension reduction and an unpunished jump to food) is not adopted, as occurs in other learning problems, also needs to be explained.

However the bimodal nature of the distribution of scores seems to be a difficulty that prevents the anxiety reduction theory from being an alternative viewpoint. Maier and Ellen previously pointed out that Farber's data (7) contained a bimodal distribution, and Farber has attempted to explain his exceptional cases by genetic strain differences and the probability of skewed distributions (7, 8). Whether or not the bimodal distributions in other sets of data (25, 27, 28, 31, 32, 37) can be reduced to artifacts of a similar nature remains a problem.

### *Can Partial Reinforcement Produce "Fixated" Responses?*

Partial reinforcement has been used to explain the fact that extinction of conditioned responses formed on the basis of 50 per cent reinforcement is slower than on the basis of 100 per cent reinforcement. If the strength of fixated responses could be reduced to partial reinforcement, it might be unnecessary to postulate a separate frustration mechanism. Since an insoluble problem also involves 50 per cent reinforcement (and 50 per cent punishment as well), and since fixations are produced more frequently in the insoluble problem test than in the soluble problems, it seems plausible to link fixations either with the frustrating effects of the insoluble problem or the partial reinforcement of the insoluble problem.

Wilcoxon (37) performed an experiment to test the applicability of the partial reinforcement concept to persistent

behavior in the jumping apparatus. He used three groups of rats, two of which were duplicates of groups used in our initial experiment (25). His Group I was like our Groups I and II which learned under standard reinforcement conditions (25). The animals were required to learn a position response in a Lashley-type jumping apparatus, and they were rewarded for a jump to the correct side and punished for a jump to the incorrect side. This situation gave the animal punishment on 50 per cent of the initial responses, and this percentage declined to 0 per cent as the animal learned the problem.

His Group III was like our Group III, since the animals were exposed to the insoluble problem situation. These animals were confronted with a problem situation in which neither card nor position was consistently rewarded. As a result animals received punishment on 50 per cent of their trials throughout the training period.

His Group II was new, and was required to learn under conditions of partial reinforcement. The rats were treated like Group I animals except that a jump to a particular side was rewarded on every *other* trial. Thus the animal was punished *every* time it jumped to the negative side and *every other* time it jumped to the positive side. The pattern of reward and punishment resulted

in animals' learning to jump to the positive side. This procedure results in animals' being punished on three out of four trials when they jump at random during the early stages of learning, and on two out of four trials when they have mastered the problem.

Table 1 shows the three kinds of groups used by Wilcoxon, the initial and final punishment-reward ratios for each, the percentage of fixated position responses obtained by Wilcoxon, and the percentage of fixated position responses obtained by Maier, Glaser, and Klee (25). For comparable groups the results are in agreement, and in neither experiment was there a problem of classifying fixated and nonfixated behaviors. The point of difference arises over the interpretation of the fact that the partial reinforcement group had 92 per cent fixated position responses whereas the insoluble problem had 58 per cent. Wilcoxon argues that partial reinforcement makes for fixations, and that the insoluble problem is a variation of a partial reinforcement situation. The experiment is not a crucial one, because according to frustration theory the greatest number of fixations should occur in the group that received the most punishment, and this is the so-called partial reinforcement group. That the frequency of fixations increases with the frequency of punishment has

TABLE 1  
FREQUENCY OF ANIMALS WITH FIXATED BEHAVIOR IN WILCOXON'S STUDY

Initial Training Situation	Initial Punishment Frequency (Random Trials)	Lowest Punishment Frequency After Learning	Wilcoxon Per Cent Rats Fixating	Maier et al. Per Cent Rats Fixating
Position reward Wilcoxon's Group I	50	0	38	30
Insoluble problem Wilcoxon's Group III	50	50	58	64
"Partial" reinforcement Wilcoxon's Group II	75	50	92	—

been reported in a number of studies (24, 27, 31, 32). Unpublished data from Takeshi Shimoyama of Tokyo show a consistent increase in the number of fixated rats with increased frequency of punishment. He used four groups (10 rats each) with punishment frequencies of 0, 30, 50, and 70 per cent, and obtained position fixations in 0, 2, 3, and 6 rats, respectively.

The partial reinforcement explanation has an advantage in that it does not make it necessary to postulate a new behavioral mechanism, whereas frustration theory does not need the concept of partial reinforcement to explain the effects of punishment and inconsistencies with reinforcement results. What is most important, however, is that the frustration mechanism is needed to account for the bimodal distribution of scores obtained in all of the groups; since partial reinforcement applies to all animals in a group, it should not cause a split.

*What Factors Influence the Number of Animals that Will Fixate a Response?*

Bimodal distributions are obtained in experiments of discrimination learning, but the proportion of animals that are unable to learn, vs. those that learn readily, varies. Chance variation and strain differences are to be expected, but a number of other causes have been experimentally demonstrated.

The following results indicate some of the relevant factors.

1. Punishment on every trial produces more fixations in rats (also more rapid learning for those that learn) than punishment on every other trial (27).

2. Sixteen days in an insoluble problem produces more fixations than eight days (24).

3. Mild electric shock given to human subjects on 75 per cent of the trials in an insoluble problem for 50 trials produces more fixations than the same ab-

solute number of shocks distributed over 150 trials (31).

4. Feeding after electric shock results in a smaller number of fixations than shock alone (7).

5. Guidance during the early stages of learning prevents fixations from forming later on when the rats are placed in an insoluble problem (20).

6. Difficult learning problems result in more fixations than easy learning problems (21, 25, 27).

7. Learned position responses that correspond to the natural preferences of the rat are less likely to become fixated than position responses that are on the side opposite the natural preference (25).

8. The use of the word "wrong" instead of electric shock, to indicate an error, produced fixations in human subjects about as frequently as electric shock to the fingers (31, 32).

9. Failure in fulfillment of expectancies seemed to be an important influence in the formation of fixations in a discrimination learning problem in rats (22).

Since punishment and experience of failure are associated with situations producing high rates of fixation, the frustration-fixation hypothesis is supported. It also seems necessary to regard the transition from problem-solving or trial-and-error behavior to stereotyped behavior to be rather sharp, or even of an all-or-nothing character. This suggests that there is a frustration threshold.

A further point of interest and possible importance is mentioned here in the hope that it may encourage other experimenters to report similar observations. This is the role of the experimenter in influencing the behavior of animals, particularly under stress. Some years ago two research assistants were working in adjacent rooms on related problems, each with three groups of



twelve or more rats from the same colony, over a period of a semester. One of them obtained the usual number of fixated position responses (over 50 per cent) in each of the successive groups with which he worked; the other was unable to obtain a single fixation. Although they compared procedures on preliminary training, methods of testing, and other general routines, they were unable to determine the reason for the differences. Motivational considerations also failed to throw light on the matter. The researcher who was unable to obtain fixations required them for his doctoral dissertation, so that his results did not correspond with his motives. However, it was discovered that he felt sorry for the rats, and this may have caused him to pet the rats between trials somewhat more than other researchers. This possible influence might be analogous to feeding after shock, which reduced the number of fixated rats in Farber's experiment (7).

#### *Can Abortive Behavior Prevent Punishment?*

The fact that animals learn to jump abortively when placed in an insoluble problem serves to reduce any punishment that a locked window produces. This method of jumping may create a condition that prevents rats from discovering that the problem is soluble, and hence interferes with learning the discrimination problem. To test this possibility Wilcoxon (37) separated the rats that learned from those that failed to learn (fixated). He found that the latter group had more abortive responses than the former, and concluded that this group difference supported the above hypothesis. This explanation could conceivably account for the bimodal phenomenon (animals that learn vs. animals that fixate). However, Wilcoxon's conclusions must be questioned because his learners came primarily from Group

I and his nonlearners from his Groups II and III. The fact that his learners and nonlearners came from different groups makes it impossible to evaluate the function of the basic variable which he purposely introduced into the experiment to test the influence of partial reinforcement. But even if this difference in abortive jumping between learners and nonlearners is obtained under comparable conditions, it is not a crucial comparison because it does not exclude the equally plausible but opposite interpretation, which is that failure to learn causes abortive jumping. In order to distinguish cause from effect, it is necessary to study how abortive jumping develops in relation to the progress made in learning.

Maier, Glaser, and Klee (25) showed that animals placed in an insoluble problem soon develop abortive jumping, which increases as the animal is retained on the insoluble problem. However, when the problem is made soluble, the abortive jumping is expressed differentially—that is to say, it declines in frequency for the trials on which the positive card is on the side to which the animal has been jumping, but increases for the trials on which the negative card is on the side to which the animal has been jumping. This, then, is the first evidence that the rat has learned to discriminate between the stimulus cards. For animals that do not show fixations, this *differential* abortive jumping begins to appear about the same time that the position response is replaced by the discrimination response; in other words, the animal now follows the positive card from side to side, and abortive jumping disappears because the animal no longer jumps to the negative card. For animals that fixate, the differential abortive jumping persists and becomes even more clear cut, since this discrimination between the stimulus cards is not followed by abandonment of the position re-



TABLE 2  
COMPARISON OF ABORTIVE JUMPING IN RATS WITH FIXATED  
AND NONFIXATED POSITION RESPONSES

	N	Per Cent Jumping Abortively	Per Cent Showing Differential Abortive Jumping	Average Days to Learn Differential Abortive Jumping
Rats with Fixated position responses (nonlearners)	75	73.3	96.4	6.6
Rats with Nonfixated position responses (learners)	29	58.6	94.1	6.1

sponse. The rat with the fixated position response fails to follow the positive card from side to side, and thus jumps abortively whenever the negative card is on the side of the fixation. The greater total of abortive jumps in rats with fixations than in rats that learn, therefore, may be an artifact, i.e., they jump abortively more often because they jump to the negative card over a longer period of time.

In a recent study (23) it was possible to analyze the data to test the hypothesis suggested by Wilcoxon. Table 2 shows that there are 73.3 per cent abortive jumpers among rats with fixations, and 58.6 per cent abortive jumpers among rats that give up the position response. Although the difference in percentages is not significant, the trend might suggest that abortive jumping hindered learning of the discrimination response. However, one may also measure learning by the *differential* abortive jumping. If jumping abortively makes the rat less prone to discriminate between the positive and negative cards, then initial abortive jumping should not lead to differential abortive jumping. However, Column 3 shows that 96.4 per cent of the animals that persist in the position response, and 94.1 per cent of the animals that adopt the discrimination response, show differential abortive jumping. These percentages are too high and too much alike to suggest that

abortive behavior can be a causal factor in the phenomenon of fixation. The rates of learning to discriminate, as measured by the number of days required to show differential abortive jumping, are shown in Column 4. These measures also are too similar (6.6 days vs. 6.1 days) to support Wilcoxon's hypothesis.

#### *Can Several Learning Concepts be Combined to Reinterpret the Michigan Findings?*

Wolpe (38) has made a systematic attempt to reduce the phenomena of abnormal fixation to learning principles by dealing with the experimental evidence on which the concept was based. He draws freely from various learning principles, and these, if not integrated, may create the dilemmas that Eglash (5, 6) has discussed in his attempts to test the adequacy of various learning concepts for interpreting the experimental findings upon which frustration theory rests. Wolpe's case will be made stronger if he can reply to Eglash and also meet the objections and questions raised below.

Wolpe accepts the anxiety reduction theory in dealing with Farber, Mowrer, and others, but at the same time he recognizes the inadequacy of the theory for dealing with the Michigan data. In order to interpret these data, he postulates that the air blast is the primary

drive and that jumping reduces the air-blast drive and hence reinforces jumping. "When jumping in a particular way has thus been repeatedly reinforced it becomes firmly established as the habitual response to the air-blast stimulus, and the more firmly it is established the weaker does the competing alternative response tendency become" (38, page 114). Thus the animal does not use its knowledge of the learning discrimination response because the air blast, not the stimulus card, is the stimulus. (It is assumed that Wolpe would deal similarly with the experiments in which electric shock was used to force a response.) This escape type of approach to discrimination learning postulates a primitive kind of S-R behavior in which one stimulus dominates the behavior to the exclusion of all others. Such stimulus dominance is not borne out by the behavior of rats in other learning situations where almost every change elicits curiosity. Five experiments on fixations seem to be in conflict with this interpretation, but some of these were too recent to be accessible to Wolpe.

Ellen (4, 17) confronted rats having fixated position responses with three stimulus cards, two of them being on the side of the animals' position fixation, the other on the side opposite the position fixation. When one of the former two was positive and the other negative, the animals tended to jump to the positive card. However, they would never select the positive card that was placed on the side opposite the fixation. Thus animals *will* jump to the positive stimulus card and avoid the negative, but only if this discriminatory behavior does not interfere with the fixated position response. Giving the animal a chance to succeed within the scope of its fixation was a type of therapy Ellen was evaluating in these experiments, and he obtained positive results of a significant

magnitude. For the purposes of this discussion, his findings demonstrate that the animals' behavior is not entirely dominated by escape in experiments with the Lashley apparatus.

Klee (14) obtained fixations in rats by using only the hunger drive to force a response. Although this was an extremely time-consuming process, and retraining had to be instituted to prevent actual starvation, fixations were produced. In this situation, a jump from the stand could not have served a reinforcement value since hunger tension was not reduced by a jump. Perhaps one might make a case that the animal was rewarded by being returned to its cage after being on the stand for many hours, but this supposition seems somewhat tenuous.

A third experiment that bears on the issue raised by Wolpe is one by Feldman (9). He has shown that if rats are required to jump to stimulus cards on some trials, and on others are presented with short pathways leading to the same cards, they will respond differently to the cards under these two conditions. Some rats had one fixation for jumping and developed another for walking when a pathway was introduced, but the same primary stimulus (electric shock to tail) was used for forcing both the walking and jumping methods for leaving the jumping stand. Interesting, too, is the fact that rats with fixated jumping responses (which Wolpe would describe as the learned responses to the primary stimulus that was applied to the jumping stand) never attempted to jump when the walking pathways were in place. Secondary stimuli influence the primary response, and this type of adjustment to a larger situation is denied by Wolpe when he speaks of secondary drives being unable to compete with the primary drive.

The fourth obstacle confronting Wolpe's basic assumption is the fact

that the number of fixations increases as the second response to be learned is increased in difficulty (21, 25). If animals merely responded to the primary drive and were not influenced by the discrimination problem with which they are confronted, the number of fixations should not be so directly influenced by the cards used.

Finally, it would seem that rats that had previously acquired fixations should develop them more quickly than others. It has been found (20, 28), however, that rats cured of fixations are less prone to develop them a second time. Although a clearly formulated explanation for this finding has not as yet been supplied, it is in conflict with Wolpe's position. It is the writer's opinion that guidance therapy increased the frustration threshold, and hence reduced the occurrence of fixations.

Experimental findings that Maier and Ellen (19) described as crucial in invalidating the anxiety-reduction theory are dealt with by Wolpe in an interesting manner. Although the basic assumption on which his viewpoint rests has been questioned above, his interpretations of the data are sufficiently ingenious to warrant consideration and comment.

a. The fact that more animals fixate the response in progress when 100 per cent punishment is used for breaking it than when 50 per cent punishment is used, is explained by saying that the 100 per cent punishment condition causes the rat to require more of an air blast to elicit a jump and hence to experience greater reinforcement of the ongoing jumping response<sup>2</sup> than does the

50 per cent punishment condition. It is not known whether 100 per cent punishment produces greater resistance than does 50 per cent punishment, but even if this were true, the argument seems to hang on tenterhooks. Would it not also follow that rats in the 100 per cent punishment condition would more readily learn to jump, having had more reinforcement? As a result they should learn sooner and therefore reduce the amount of air blast. Reducing the air-blast exposure would decrease the reinforcement of the 100 per cent group, thus allowing the 50 per cent punishment group to catch up.

b. Rats in the 100 per cent punishment group that give up their position responses do so on fewer trials than rats that are in the 50 per cent punishment group. This is explained by saying that only rats that learn quickly avoid developing the fixations described above. This is in general agreement with the explanation Maier and Klee used (27), except that they indicated that pressure from behind (shock) increases motivation until the pressure reaches the point where it is frustrating to animals. Animals would differ at which point pressure ceases to be motivating and becomes frustrating.

c. To explain the bimodal distribution of scores, Wolpe uses the principle of response oscillation. When the response strength of an alternative response is strong enough, it may occasionally be expressed, and the fixation may be broken. In other animals the alternative response does not happen to occur while the relative habit strengths are right. This apparently means that if an animal tries out another response, it can discover the virtues of this response and adopt it if the strength of the primary response is not too great. However,

normal range. The most frequent criterion is 200 trials of no change when a correct response is clearly available.

<sup>2</sup> Wolpe describes the ongoing response as a fixation, but he is actually referring to position stereotypes and position habits. This error is beside the point for the purpose of the present discussion, but is mentioned in order to correct a frequent misconception. According to frustration theory, not every persistent response is an abnormal fixation, but only those that have been shown to fall outside the

this concept seems to overlook the compulsive character of the fixated response.

d. Wolpe deals with some of the evidence that has been used to make a case for the compulsive behavior. He points out that an animal does not jump to an open window when the opportunity is given because the response to the air blast is so strong that it completely dominates the animal's behavior. This is the kind of attention one may expect in panic, but it seems to be a rather extreme assumption to state that a rat in a problem-solving situation can be so completely controlled by a stimulus.

e. Failure of trial-and-error behavior to occur when a fixated response is punished is given as evidence that a fixated response is a strong habit. He points out that the rigidity of the position response depends upon the fact that reinforcement occurs on all occasions. Thus the compulsive character of behavior is on a continuum with strong habits.

f. The fact that guidance may effectively break a fixation in a few trials is not regarded as an unusual phenomenon by Wolpe. Guidance, he points out, makes an alternative response possible and so the very strong habit is dropped for another. Why do relative habit strengths sometimes change so quickly, while at other times the outstanding characteristic of competing responses is the complete domination of a single habit? Are the explanations for points e and f parts of one learning theory, or do they represent principles from competing learning theories? Detailed analysis of the data on the effects of guidance demonstrate that an animal does not adopt an alternative response immediately; rather, the effect of guidance is that it causes the rat to give up the fixated response. Only later is the new response learned. Maier and Klee (28) have shown that guided animals show a gap between giving up the fixated response and adopting the new response,

whereas nonguided animals do not give up the response in progress until an alternative is practically learned.

Wolpe does not claim that frustration theory has been disproven by his discussion. Rather he feels the theory is unnecessary because an explanation in terms of learning is possible.

*Can the Difference Between Fixations and Habits be Reduced to Two Kinds of Learning?*

In 1935 Maier and Schneirla treated conditioning and trial-and-error learning as qualitatively different forms of learning, and discussed the topics in separate chapters of their book (29). Later (30) they pointed out how conflicting evidence in learning data could be reconciled by making a qualitative distinction between trial-and-error learning, which always requires reinforcement in the form of reward (or punishment), and Pavlovian conditioning, which requires reinforcement in the form of the presentation of the unconditioned stimulus. Confusion arose because in some instances the same object (food) could serve either as a stimulus for salivation or as a reward for the same response. Skinner (35) later found a similar distinction (respondent vs. operant conditioning) essential to his organization of data, while Mowrer (33) raised the same issue when he postulated the dual nature of learning: conditioning of autonomic responses (such as fear reactions) and instrumental (trial-and-error) learning. Recently Solomon and Wynne (36) have used this dual nature of learning as a point of departure for explaining the rapid acquisition and unusual persistence of avoidance responses. They believe that different sets of learning concepts must be used when pain-fear stimulation is intense enough to cause trauma, and when dealing with mild avoidance-producing stimuli.

In introducing stimulus intensity

(classified as trauma-producing) as a learning variable, Solomon and Wynne must assume that there is a threshold that has to be exceeded to obtain the change, and they indicate that a state of anxiety, with the resultant autonomic responses and feedback, is produced with such stimuli. This viewpoint is in the direction of the frustration theory, but it fails to take the complete step of postulating a dual mechanism. It can explain why mild and intense punishment produce different reactions, and, if the threshold concept is clarified, it can explain bimodal distributions. The theory is too limited, however, to incorporate regressive and aggressive behaviors, and it does not escape the onus of postulating a qualitative distinction, since behavior developed in a state of trauma differs from that learned under nonanxiety-producing conditions.

There is considerable evidence which indicates that all learning data cannot be incorporated under one set of principles, but the nature of the essential difference is controversial. The fact that the Pavlovian type of conditioning experiments often utilized autonomic responses may have caused Mowrer, as well as Solomon and Wynne, to make the separation along physiological lines. However, a conditioned eyewink is an example of strict Pavlovian conditioning, but it is not associated with anxiety. It would also be inappropriate to classify eyelid conditioning with instrumental learning.

#### EXTENSIONS OF FRUSTRATION THEORY

##### *Physiological Basis for Frustration Mechanism*

In describing the frustration mechanism, emphasis has been given to differentiating it from the motivation process at the behavior level. Is there a physiological justification for postulat-

ing a second type of intervening variable? The one that naturally comes to mind is to associate frustration with the autonomic nervous system. The concept of threshold is associated with emotional behavior and the arousal of the autonomic nervous system. Emotional behavior is often described as irrational and uncontrollable, and placed in contrast with rational behavior. Although frustration and emotion in many respects are overlapping concepts, they are by no means equivalent terms. For example, the emotions include love and pleasant feeling tones as well as fear, hate, and unpleasant feeling tones; whereas frustration is usually used as a more specific and less inclusive term. For the present it seems appropriate to link rage and terror with frustration, but to exclude the pleasant emotions and mild emotional feeling tones. Since frustration theory clearly separates hate and intense fear states from emotions of love, it becomes imperative to find differing autonomic reactions during states of love and hate if support for this physiological mechanism is to be obtained.

Jost has analyzed a variable in polygraphic measurements that may throw new light on this problem (11). He emphasizes the *patterning* of the various polygraph indices rather than the individual measures of heart rate, blood pressure, galvanic skin resistance, and muscle tension, and describes some interesting individual differences in the manner of reacting to conflict. Using stressful and frustrating situations, Jost, Ruilman, Hill, and Gulo (12) were able to differentiate between the autonomic reactions of hypertensive patients and a normal control group. Jurko, Jost, and Hill (13) also obtained significant differences in physiological measures between nonpatient, psychoneurotic, and early paranoid schizophrenic groups. The authors present evidence favoring differences in patterns of discharge



between the groups, and suggest the concept of levels of energy discharge. Psychoneurotics tend to show hyperadaptation, whereas schizophrenics show disadaptation. They used the Rosenzweig Picture Frustration Test to produce the emotional changes. These studies are suggestive in that they reveal that persons suffering from emotional problems show physiologically different reactions to frustration than normals. In an earlier study, Sherman and Jost (11) had found physiological measures to be sufficiently unique to serve a diagnostic purpose.

Jost (11), as well as Solomon and Wynne (36), stresses the notion that the autonomic functions may influence higher neural centers by some feedback mechanism. This type of control over higher centers would seem to be consistent with frustration theory, and studies along this line should lead to further development and refinement of the theory.

In relating the autonomic nervous system and the frustration process one cannot assume that autonomic functions are absent when a state of frustration does not exist. Rather, the experimental question becomes one of determining whether a level, or unique patterning, of autonomic functioning can be associated with frustration. Certainly pleasant and unpleasant feeling states can occur without assuming frustration. At what level does anger change to rage, fear to terror or panic, and choice behavior to compulsion? Frustration theory would demand some kind of sharp transition, either in the form of a dropping out of voluntary control mechanisms or a sheer dominance of autonomic processes as provocation exceeded a certain point. The possibility of an intermediate condition, such as conflict between the motivation and frustration processes, needs exploration, and it may be the kind of

mechanism that underlies the state of anxiety.

Another physiological mechanism that conceivably may be associated with the frustration process is suggested by the researches of Coghill (2). It will be recalled that he found behavior changes in the developing salamander to run parallel with neural growth; new behaviors appeared rather suddenly as essential neural links were formed. According to Coghill's work, the gross bodily functions develop first, and require only a limited number of cells in a chain of neurons. For example, the limb responses of the salamander occur in conjunction with waves of contraction that move down the side of the body. A single stimulus thus releases this total behavior. Later, as the limbs develop further, additional neurons are grown so that each muscle group has its own specific neurons and therefore controls a smaller amount of behavior. Eventually each movement has its own controls. The process of development therefore is from gross or total behavior to specific or small units of behavior. This view of the developmental process, known as *individuation*, is in contrast to the view that treats the reflex as the unit of behavior.

According to Coghill's analysis, it follows that each muscle group may be activated either as part of the larger pattern or as a separate response. Specific or individuated behavior depends upon the acquisition of additional neural mechanisms which are superimposed upon the more primitive mechanisms controlling the grosser patterns. However, the newer responses can make their appearance only if there is a successful inhibition of the older patterns. Thus mature behavior is both a process of excitation of the more recently developed pathways and a process of inhibition of the older pathways.

Up to this point the anatomical evi-



dence is fairly rich, and the speculation is supported by the evidence. The next step requires merely that we assume that frustration removes inhibition and permits the grosser responses to take over in much the same way that removal of the cortex removes spinal inhibitions. Since frustrated responses are more global and less subject to cortical control, the required assumptions are not unreasonable. Certainly the general notion seems worthy of further refinement.

If the evidence for qualitative difference in behavior is sound, it becomes important to locate the physiological basis on which it rests. Research along this line could do much to increase our knowledge of the control of behavior.

In attempting to relate frustration-instigated behavior both to the autonomic nervous system and to grosser segments of unindividuated behavior, a relationship between emotion and primitive mass behavior is implied. This connection is not out of line with basic facts, because emotions are integrated at the thalamic level, they are general or even total responses, and they are pretty much beyond voluntary control. Since the cortical mechanisms are new and superimposed upon the phylogenetically older subcortical structures, it is understandable if their influence is confined to ideation and motivated or rational behavior.

#### *The Function of Catharsis*

The benefits attributed to catharsis vary and they perhaps are accepted more readily by the clinician than by the theoretician. From the point of view of frustration theory, catharsis should have a therapeutic value of two sorts.

When in a state of frustration, during which the autonomic nervous system sends sensory impulses (feedback) to the central nervous system, an individ-

ual's behavior becomes the end of a sequence. The behavior is the organism's response to the frustrating condition as it exists at that moment. Since the frustrated behavior is the final link in the causal chain, anything that facilitates its expression tends to end the cycle. Behaviors released along harmless channels (where there is no retaliation) most effectively remove the stimulus conditions setting off the state of frustration, whereas forms of physical violence that are met with superior force are least beneficial because they set up a new cycle of frustration. The unit of behavior controlled by the frustration process is in contrast with that controlled by the motivation process, because the latter requires a goal or need reduction to terminate its cycle, and behavior expressed serves as a means to this end.

The second way in which catharsis serves a therapeutic function according to frustration theory is through a clarification of the problem. As long as a state of frustration exists, the autonomic nervous system sends distracting sensory impulses to the brain. These are experienced as feelings (in accordance with the James-Lange theory) and tend to confuse the cognitive picture of the problem situation. Clarification of the perception of the actual problem should occur, both because verbal expression would tend to unscramble the two conflicting sources of data (externally vs. internally aroused sensations) and because general expression should reduce the extent of the confusing feeling data.

#### *The Unconscious and the Logic of Feeling*

Freudian theory describes the need to bring the unconscious mind to consciousness in overcoming compulsions. Many forms of conflict are therapeutically treated as though they were the result of contradictory influences of the

conscious and the unconscious minds. In brief, two different kinds of influences on behavior are postulated. Frustration theory sets up a similar dichotomy: the logic of ideas (i.e., of the mind) and the logic of feeling (i.e., of the viscera). Since these logics are qualitatively different, an aspect of therapy in frustration theory depends upon removing the conflicts between thinking and feeling.

A little further speculation brings these two sets of dichotomies close together. Feelings are vague and difficult to describe, but most compelling. Was it not natural for a theorist, writing at the time when psychology was dominated by philosophy, to describe these confusing and yet compelling forces as an unconscious mind? In the writer's thinking this simple step of substituting feeling for the unconscious mind has given him a greater appreciation of the original psychoanalytic concepts, and at the same time it clarifies the place of emotion in psychological conflicts.

#### *Preoccupation with the Self and Therapy*

If the state of frustration is characterized by autonomic feedback, it follows that feelings will dominate consciousness at such times; as a result, perceptions and thoughts will have subjective rather than objective reference. The frustrated person, therefore, must be unusually aware of a mass of feelings or sensations of internal origin. When these sensations are excessive, they would occupy his attention and consequently make him less aware of the external world as well as insensitive to the feelings of other persons. His social behavior would appear selfish and inconsiderate because he would be preoccupied with his own feelings and his relationship to other persons, rather than concerned with problems involving relationships to external objects. He

would be a good observer of how another person's actions affect him, but quite unaware of the way his own behaviors affect the feelings of others. Such a person would be preoccupied with himself and the nature of the self.

Viewed in this manner, excessive preoccupation with the self would tend to make a person ineffective in group endeavors. A person who contributed to a discussion only for the purpose of gaining something for himself or even for the purpose of making an impression on others would not be functioning at his best in assisting the solving of external problems. Behaviors such as self-aggrandizement and self-effacement also seem to be by-products of an excessive awareness of the self, and serve only incidental purposes.

Even the solving of human-relations problems would not be promoted by a person whose perceptual organization was dominated by sensations of internal origin. Any sensations that tended to increase the awareness of the self would be inclined to cause a person to see himself as a central figure in a social situation. This kind of social outlook would readily lead to misunderstandings and conflicts. Good human relations would seem to require sensitivity for the feelings of others.

The condition of paranoia illustrates how the external world may be distorted by being viewed against a background of feeling. The self seems to be the center of the universe, and all events and actions of others become related to the self. Thus the meanings and interpretations of events and actions are influenced by the observer's role in the universe and by the way he feels. Although the paranoid personality is aware of his external world, he is not able to separate the sensations of external origin from the mass of feelings created by his anxieties.

The distracting role of visceral sensa-

tions becomes even more apparent in the condition of schizophrenia. In describing the nature of the schizophrenic process, Jenkins (10) has emphasized both the early frustration and the lack of interest in external things. He describes the self-consciousness of patients as painful and maladaptive. The schizophrenic becomes insensitive to minor social cues (intonations, gestures, and so forth), and more and more his consciousness centers "in processes which are the very essence of the self" (10, p. 9). Thus the schizophrenic does not seem to retreat from the external world; rather, his inner world takes over. It appears that the withdrawal behavior of schizophrenics may not be a form of escape from the external world, but one of being engulfed by his internal world. This distinction would seem to have basic therapeutic implications.

The fact that the literature on clinical psychology deals considerably with the problem of the self, while the literature on normal personality shows little concern with the problem, supports the hypothesis described above. If this idea is pursued a step farther, it suggests the notion that good adjustment is a process of being aware primarily of external stimulation. Therapy, accordingly, would become the process of reducing the awareness of the self as well as the amount of subjective sensations, and increasing the awareness of external stimulation, including the behavior of other persons. Human relations programs, utilizing discussion and role-playing procedures, are examples of training methods that have been used to increase the sensitivity of an individual to the feelings of others. It is possible that too much concern with the self may be detrimental, and that the type of interest a poorly adjusted person portrays may delay his adjustment.

Both therapy and human-relations training have as their objective the

building of constructive attitudes and improved communication between client or trainee and other persons. Training in human relations has profited much from clinical psychology and it is now possible that clinical psychology may gain something from the experiences in human relations training where group methods and concern for the other person are emphasized in the methodology (18).

### *Sensory Deprivation and the Self*

There are two ways in which a person might conceivably become preoccupied with his internal sensations, if we assume that domination is a relative matter. The examples cited above are taken from abnormal psychology and illustrate the condition of being trapped or overwhelmed by an overabundance of sensations of internal origin. The second possible type of preoccupation with internal stimulation could result from a sharp reduction in external stimulation.

The researches associated with extreme sensory deprivation indicate that when external stimulation is greatly reduced, the content of consciousness becomes startlingly different and strange to the person so treated. Much of this research is, as yet, unpublished and it is difficult to anticipate how the data will add up, but there is a strong hint that abnormal emotional experiences and anxiety states are produced in this manner. For example, Bexton, Heron, and Scott (1) reported various kinds of hallucinations and dream-like imagery when their subjects were placed on a comfortable bed in a lighted, soundproof cubicle for an extended period. To reduce sensations further, subjects wore translucent glasses, gloves, cardboard cuffs that extended beyond the finger tips, and a U-shaped rubber pillow which restricted movement.

Subjects found the experience unpleasant; they showed emotional lability, were easily upset, and irritability increased as the experimental period continued. In addition to visual, auditory, tactual, and kinaesthetic hallucinations, some subjects described strange experiences. Two subjects reported that they felt as if they had two bodies; in one case the bodies were side by side and in the other, they overlapped. Other subjects reported feelings of "otherness" and bodily "strangeness." Others reported their heads or minds as being detached from the body.

The authors conclude that their findings represent evidence of a kind of a dependence a person has upon the sensory input that previously had not been recognized. It would seem to require but little stretch of the imagination to interpret the experiences reported as containing a description of an exaggerated awareness of the self, and to regard the hallucinations produced under sensory deprivation to be similar to those reported by mental patients.

The fact that sensory deprivation tends to produce a wide variety of new and strange experiences for subjects, which also seem to be highly unpleasant (1), has led to a good deal of speculation. Since the findings came as a distinct surprise to psychologists generally, it is apparent that our present knowledge of human experience does not permit us to predict the effects of extreme changes. In other words, accepted theory does not permit realistic speculation that wanders too far from previously explored territory.

If, however, the phenomena are viewed from the frame of reference of frustration theory, no new assumptions are needed. Rather the findings supplement the interpretation that this theory places upon mental illness, and even lead one to entertain seriously the consideration of a form of therapy for cer-

tain conditions of schizophrenia that is either new or that previously lacked conceptual support. This is the use of intense external stimulation to offset the excessive internal stimulation. It is possible that music therapy served this distracting purpose for some patients, but the general idea suggested by these experiments goes beyond this point. Among the research approaches that come to mind are the use of amplification devices in speaking to patients; making radical changes in the environments of patients, such as the use of bright lights and sharp contrasts; and varied methods designed to increase activity. It is conceivable that the intensely stimulating environment would serve to release withdrawn patients from the domination of stimulation of internal origin and thereby make them more accessible to psychological therapy. Drugs that depress visceral functions should serve a similar purpose.

#### SUMMARY AND CONCLUSIONS

This paper has two major objectives: (a) to clarify vague or misunderstood conceptions of the author's theory of frustration, and (b) to extend the theory and its implications.

Nine questions are discussed in an attempt to achieve the first objective. The greatest source of confusion and difference in viewpoint is caused by the postulation of a qualitative distinction between frustration-instigated and motivation-controlled behavior. This distinction requires a re-examination of psychological causation and the part played by learning in behavior modification, but each alternative approach seems to satisfy one portion of the evidence at the expense of another portion. The systematic evaluation of the facts regarding the varied issues raised by the questions should clarify problems for future research.

Other issues, around which questions conceivably could be raised, have to do with the *threshold* concept implied in frustration theory and the concept of *availability* as the selective mechanism in frustration. However, these seem to be secondary issues at the present time because they depend upon the initial acceptance of the qualitative distinction. It is hoped that the primary assumption of a qualitative distinction will not continue to prevent the investigation of the many new problems that are suggested by the assumption.

The new developments in frustration theory deal with the implications of two sources of sensory data: those of internal origin and those of external origin. The discussion includes (a) suggestions for a physiological basis for two qualitatively different types of behavior, (b) an analysis of the therapeutic value of catharsis, (c) a comparison between the logic of feeling and the functions of the unconscious, (d) the concept of the self in the light of frustration theory and its implications for therapy, and (e) an interpretation of the results of sensory deprivation in the light of the two sources of sensory data.

## REFERENCES

1. BEXTON, W. H., HERON, W., & SCOTT, T. H. Effects of decreased variation in the sensory environment. *Canad. J. Psychol.*, 1954, 8, 70-76.
2. COGHILL, G. E. *Anatomy and the problem of behavior*. London: Cambridge Univ. Press, 1929.
3. DOLLARD, J., & MILLER, N. E. *Personality and psychotherapy*. New York: McGraw-Hill, 1950.
4. ELLEN, P. The compulsive nature of abnormal fixations. *J. comp. physiol. Psychol.*, 1956, 49, 309-317.
5. EGLASH, A. Perception, association and reasoning in animal fixations. *Psychol. Rev.*, 1951, 58, 424-434.
6. EGLASH, A. The dilemma of fear as a motivating force. *Psychol. Rev.*, 1952, 59, 376-379.
7. FARBER, I. E. Response fixation under anxiety and non-anxiety conditions. *J. exp. Psychol.*, 1948, 38, 111-131.
8. FARBER, I. E. Anxiety as a drive state. In M. R. Jones (Ed.), *Nebraska symposium on motivation*. Lincoln, Nebraska: Nebraska Univ. Press, 1954. Pp. 1-55.
9. FELDMAN, R. S. The specificity of the fixated response in the rat. *J. comp. physiol. Psychol.*, 1953, 46, 487-492.
10. JENKINS, R. L. Nature of the schizophrenic process. *Arch. Neurol. Psychiat.*, 1950, 64, 243-262.
11. JOST, H. *The use of polygraphic techniques in psychophysiological research and clinical psychology*. Chicago: Associated Research, Inc., 1953. Pp. 11.
12. JOST, H., RUILMANN, C. J., HILL, T. S., & GULO, M. J. Studies in hypertension: I. techniques and control data; II. central and autonomic nervous system reactions of hypersensitive individuals to simple physical and psychological stress. *J. nerv. ment. Dis.*, 1952, 115, 35-48; 152-162.
13. JURKO, M., JOST, H., & HILL, T. S. Pathology of the energy system: an experimental-clinical study of physiological adaptive capacities in a non-patient, a psychoneurotic and an early paranoid schizophrenic group. *J. Psychol.*, 1952, 33, 183-198.
14. KLEE, J. B. The relation of frustration and motivation to the production of abnormal fixations in the rat. *Psychol. Monogr.*, 1944, 56, No. 4 (Whole No. 257).
15. KLEEMEIER, R. W. Fixation and regression in the rat. *Psychol. Monogr.*, 1942, 54, No. 4 (Whole No. 246).
16. MAIER, N. R. F. *Frustration: the study of behavior without a goal*. New York: McGraw-Hill, 1949.
17. MAIER, N. R. F. The premature crystallization of learning theory. In *The Kentucky symposium: Learning theory, personality theory, and clinical research*. New York: Wiley, 1954. Pp. 54-65.
18. MAIER, N. R. F. *Principles of human relations: applications to management*. New York: Wiley, 1952.
19. MAIER, N. R. F., & ELLEN, P. Can the anxiety-reduction theory explain abnormal fixations? *Psychol. Rev.*, 1951, 58, 435-445.
20. MAIER, N. R. F., & ELLEN, P. Studies of abnormal behavior in the rat: the prophylactic effects of "guidance" in reduc-

- ing rigid behavior. *J. abnorm. soc. Psychol.*, 1952, **47**, 109-116.
21. MAIER, N. R. F., & ELLEN, P. Reinforcement vs. consistency of effect in habit modification. *J. comp. physiol. Psychol.*, 1954, **47**, 364-369.
  22. MAIER, N. R. F., & ELLEN, P. The effect of three reinforcement patterns on position stereotypes. *Amer. J. Psychol.*, 1955, **68**, 83-95.
  23. MAIER, N. R. F., & ELLEN, P. Studies of abnormal behavior in the rat: XXIV. Position habits, position stereotypes and abortive behavior. *J. genet. Psychol.*, 1956, **89**, 35-49.
  24. MAIER, N. R. F., & FELDMAN, R. S. Studies of abnormal behavior in the rat: XXII. Relationship between strength of fixation and duration of frustration. *J. comp. physiol. Psychol.*, 1948, **41**, 348-363.
  25. MAIER, N. R. F., GLASER, N. M., & KLEE, J. B. Studies of abnormal behavior in the rat: III. The development of behavior fixations through frustration. *J. exp. Psychol.*, 1940, **26**, 521-546.
  26. MAIER, N. R. F., & KLEE, J. B. Studies of abnormal behavior in the rat: VII. The permanent nature of abnormal fixations and their relation to convulsive tendencies. *J. exp. Psychol.*, 1941, **29**, 380-389.
  27. MAIER, N. R. F., & KLEE, J. B. Studies of abnormal behavior in the rat: XII. The pattern of punishment and its relation to abnormal fixations. *J. exp. Psychol.*, 1943, **32**, 377-398.
  28. MAIER, N. R. F., & KLEE, J. B. Studies of abnormal behavior in the rat: XVII. Guidance versus trial and error in the alteration of habits and fixations. *J. Psychol.*, 1945, **19**, 133-163.
  29. MAIER, N. R. F., & SCHNEIERLA, T. C. *Principles of animal psychology*. New York: McGraw-Hill, 1935.
  30. MAIER, N. R. F., & SCHNEIERLA, T. C. Mechanisms in conditioning. *Psychol. Rev.*, 1942, **49**, 117-134.
  31. MARQUART, D. I. The pattern of punishment and its relation to abnormal fixation in adult human subjects. *J. gen. Psychol.*, 1948, **39**, 107-144.
  32. MARQUART, D. I., & ARNOLD, L. P. A study of the frustration of human adults. *J. gen. Psychol.*, 1952, **47**, 43-63.
  33. MOWRER, O. H. On the dual nature of learning—a reinterpretation of "conditioning" and "problem-solving." *Harv. educ. Rev.*, Spring, 1947, 102-148.
  34. SHERMAN, M., & JOST, H. Diagnosis of juvenile psychoses. *Amer. J. Dis. Child.*, 1943, **65**, 586-592.
  35. SKINNER, B. F. *The behavior of organisms*. New York: Appleton-Century, 1938.
  36. SOLOMON, R. L., & WYNNE, L. C. Traumatic avoidance learning: the principles of anxiety conservation and partial irreversibility. *Psychol. Rev.*, 1954, **61**, 353-385.
  37. WILCOXON, H. C. "Abnormal fixation" and learning. *J. exp. Psychol.*, 1952, **44**, 324-333.
  38. WOLPE, J. Learning theory and "abnormal fixations." *Psychol. Rev.*, 1953, **60**, 111-116.

(Received January 30, 1956)



## THE INADEQUACY OF THE HULLIAN DERIVATIONS OF REASONING AND LATENT LEARNING

J. A. DEUTSCH

*Institute of Experimental Psychology, University of Oxford*

Many of the Hullian derivations concerning reasoning and latent learning which use the antedating goal reaction (or the anticipatory goal response) are not logically sound. The oversight is of the following nature: In the usual instance an antedating goal reaction ( $r_0$ ) will undergo extinction if the goal reaction of which it is a fraction does not occur and provide reinforcement. For instance, an antedating food response will be weakened if the animal is no longer rewarded in the goal box. But suppose that instead of a mere withdrawal of food, another reward occurs instead. The antedating response to food, having been evoked, will now continue to be reinforced. In the derivations to be examined, Hull appears to overlook that the antedating response would continue to be reinforced to the "wrong" goal if the goal reward changed, and so extends the case where reward is withdrawn to the case where it is substituted. This extension is not warranted by the premises used in the derivations, nor, as far as the writer can see, by anything in Hull's system. It seems, however, that Hullian theory, if it is to be persuasive, must be more explicit in these derivations—either by way of showing how its present suppositions may be used or by introducing fresh assumptions.

The only way that the writer has been able to explain the consistent illicit reasoning, when anticipatory goal response is used, is to assume that, having explained the formation of anticipations and their frustrations

in the simple case, Hull assumed that he could explain more complex cases of anticipatory phenomena. Confirmation and reinforcement are coextensive when an anticipatory goal response of food is subsequently reinforced by food in the goal box. Similarly frustration and extinction are coextensive when an anticipatory goal response is not followed by reinforcement. But confirmation and reinforcement are not coextensive when an animal which is both hungry and thirsty makes an anticipatory response characteristic of eating, but is rewarded by water instead. Here the anticipatory response is reinforced, but the anticipation of food, in ordinary language, is frustrated. The simplest way to explain the Hullian error is to assume that the everyday meaning of anticipation has slipped into Hull's thinking where only the technical connotation is warranted. As this assumption brings out the unitary nature of the various difficulties in the proofs to be discussed, and brings out the gap between premises and conclusion more vividly, it is used by the writer as an expository device. The validity of this psychological interpretation of the causation of the logical inadequacy of the proofs does not affect the validity of the claim that such a logical inadequacy in fact exists. Neither does this interpretation amount to a claim that reinforcement and confirmation are necessarily different and irreducible in general, but only if we use the premises which Hull used in his derivations.

It was one of the axioms of Hull's

thought that behavior at any time is always the product of stimulation at an instantly preceding moment. Hence if such stimulation is not observable it is assumed to exist. Such a situation arises in the case of behavior which is called motivated. When an animal is hungry its behavior undergoes a change. To Hull such a change must be the product of some stimulation. Accordingly Hull postulates that there is a drive stimulus ( $S_D$ ) operating. Further, he assumes that this inaccessible stimulus will be correlated with other changes in the animal's behavior, in the same way as the observable events in its class. For instance, it will tend to evoke any response which it preceded when a reinforcement has occurred (5, Postulate III, p. 5).

There is another type of behavior which appears to be determined not by stimulation impinging at the time, but by a distant goal. Here Hull suggests that a part of the stimulation characteristic of the goal is acting to determine the correct response. This raises a further problem. How is a part of the goal stimulation displaced? Hull again solves this by extrapolating from observed relationships. "Consider an organism which is presented with a sequence or chain of external stimuli,  $S_1S_2S_3S_4$ , and  $S_G$ , and which makes a sequence of responses  $R_1R_2R_3R_4$  and  $R_G$ , where  $S_G$  is the food stimulus and  $R_G$  is the consummatory response, e.g., that of eating. The organism is assumed to be hungry, so that  $S_{Dh}$  will accompany the  $R_G$ , or eating response. The preceding considerations show that  $R_G$  may be reinforced to the persisting  $S_D$  and to the rather differently persisting traces of  $S_1S_2S_3S_4$  and  $S_G$ ." This is Hull's solution to the question: How is a part of the goal stimulation displaced? How, when it has been displaced, does

it produce a performance leading to the goal?

It follows that on a repetition of this sequence there will be a tendency for  $S_D$ , together with the traces of  $S_1, S_2, S_3$ , and so on, to evoke  $R_G$  at the outset of the sequence and more or less continuously throughout it except in so far as there may be a conflict between  $R_G$  and the necessary instrumental movements of the sequence, such as  $R_1, R_2, R_3$ , and  $R_4$ . Presumably in any such situation the instrumental acts would dominate the conflicting portion of the antedating generalized act, permitting the non-conflicting or fractional portion to persist in a covert form (5, p. 125).

This fractional portion he calls  $r_G$ , which generates  $s_G$ . It occurs as a part of a chain of stimuli and responses which are reinforced, that is, close to the rapid diminution of the drive stimulus or to neutral stimuli that have themselves occurred close to such a diminution. The antedating goal reaction  $r_G$  and the fractional goal stimulus  $s_G$  thus obey the same rules as other stimuli and responses.

Hull was partly led into this extrapolation from the observed by his belief that this extension reduced the phenomena of insight and reasoning to the empirical relations to be found in simple forms of learning. "The writer is inclined to the view that the principles of association between stimuli and responses, particularly as revealed in modern conditioned-reaction experiments, offer a possibility of explanation in a manner which Maier's analysis failed to take into consideration" (4, p. 221). Hull has been criticized on the grounds that his concept of the antedating reaction is vague. However, to say that a concept is vague neither destroys its usefulness nor gives grounds for its rejection. If by making such an assumption we drastically simplify the picture of animal behavior, we have some reason to believe that we are at least working on the right lines. This

makes it likely that greater precision will emerge.

### Criticism

Though their vagueness gives them a great deal of elasticity, Hull's assumptions, if consistently applied, do not in fact enable us to predict the phenomena of insight and reasoning, or even many forms of latent learning. This may be seen by examining some of Hull's explanations of these abilities.

Hull, in one of his papers (4), attempts to reduce reasoning in the rat to principles of conditioning. He imagines a problem bearing some vague resemblance to Maier's, and one which we have no evidence that the rat can solve. He then puts forward an analysis of the rat's successful solution.

There are four distinctive boxes R, U, X, and H; as shown in Fig. 1, R is connected to all other boxes and U is also connected to X. The animal is first to be trained hungry to go from R and U, separately, to X, where it is fed. Then it is to be trained to go from R to U and from R to H, as distinct habits for a reward of water. When the tendencies to go to U and H from R are equal, when the animal is thirsty, it is to be made hungry and placed on R with the path from R to X blocked. It is assumed that the animal would go to U and thence to X.

Briefly, Hull says that when the animal has been trained to go from U to X when hungry, hunger stimuli will tend to evoke the anticipatory reaction  $r_u$ . Further, when the animal is run thirsty from R to U, the platform R will tend similarly to elicit  $r_u$ , "brought forward to the beginning of the series presumably through its association with  $S_D$  or through the action of trace reactions while in their early stages (or both)" (5). Then when the animal is placed on R when hungry (and path RX blocked),  $r_U$  will be

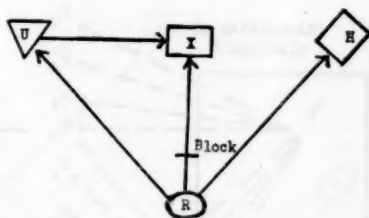


FIG. 1. A diagram of the maze in Hull's adaptation of Maier's reasoning experiment.

evoked both by the  $S_D$  of hunger, and the stimulation at R, whereas all the other anticipatory responses (there are eleven on Hull's diagram) will have been elicited only by the stimulation at R ( $S_R$ ). This will produce  $s_u$ , which in conjunction with  $S_R$  will produce the response  $R_R$ , which is a part of the response chain from R to U. The relevant part of Hull's diagram is reproduced.

Hull argues that  $R_E$ , which is a part of the response chain from R to H, will not be similarly evoked, because it "has only one excitatory tendency," that originating from  $S_R$ . This is not entirely true, for  $S_R$ , according to Hull, arouses all the anticipatory components of the responses leading to H, which in turn evoke their appropriate stimuli. These, having occurred regularly in conjunction with  $R_E$  before the animal was rewarded at  $R_H$ , would tend to evoke  $R_E$ . However,  $s_u$  is stronger than these other stimuli,  $r_u$  being the product not only of  $S_R$  but also of the  $S_D$  of hunger. We may thus, with Hull, concentrate merely on  $r_u$  and  $s_u$ , and examine not solely the critical trials but also the training series to which he has not cared to apply his premises. Suppose (as Hull does) that the animal has already learned the habit segment U to X and R to U. Then  $r_u$  and  $s_u$  will occur, evoked by  $S_R$  and  $S_D$ , when the animal is placed thirsty on platform

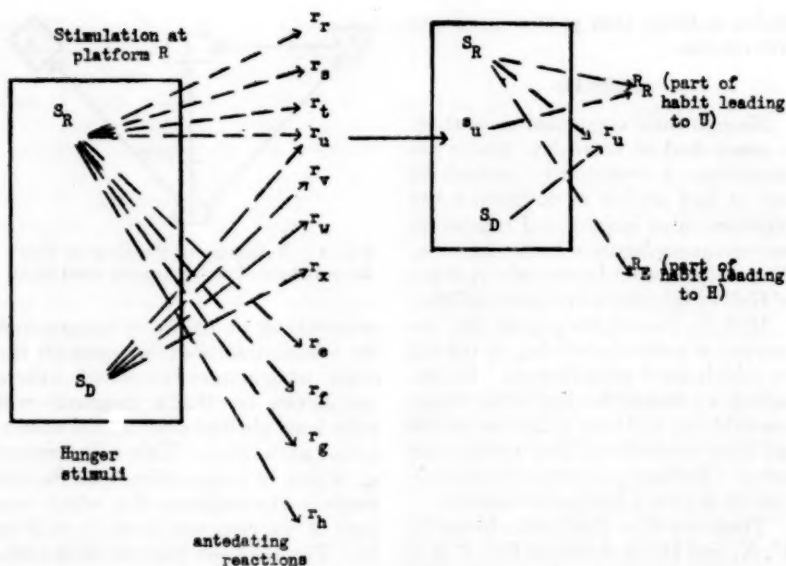


FIG. 2. Annotated version of a part of Hull's diagram. The diagram represents the anticipatory goal responses which according to Hull would all be aroused by  $s_R$  and  $s_D$  of hunger.  $R_R$  is the correct response and is evoked by  $s_R$ . The criticism is that if  $r_u$  is aroused by  $s_R$ , then  $s_R$  should have equally been conditioned to  $R_E$  during training.

R and is then run to be rewarded at H;  $r_u$  and  $s_u$  will thus equally become a part of the initial part of the habit segment leading to H. In fact  $r_u$  will be present at R when the animal is being run to H from the very first trial. This cannot be the case when it first runs from R to U. Having been more often rewarded, the association should therefore be stronger between  $s_u$  and  $R_E$  than between  $s_u$  and  $R_R$ . Thus if Hull predicts anything at all, he should predict that when the animal is placed hungry on R, it should run to H—a stupid performance.

This objection applies equally if  $r_x$  has been conditioned back to occur at R, as in Osgood's account (7). It seems that, on a strict interpretation of Hull, any antedating response, if it has been evoked, will become a part

of any behavior segment whatever reward it antedates. On the other hand, Hull, when he knows that the anticipation in ordinary parlance would be incorrect, tacitly applies the common-sense idea that an anticipation is reinforced only when it is confirmed, not rewarded. An interesting example is to be found in his last book, where he attempts to deduce another imaginary Maier-type experiment. Here both the strict rule and common-sense version are applied at different points. If either version were consistently to be used throughout, the deduction could not be made.

To this example we shall now turn (5, p. 310). The animal is first trained to run from J to L (as in Fig. 3) for a very large food reward. Then it learns on a separate occasion to run from H to J for a small food reward,

and subsequently from H to N for a similarly small food reward. Then it is placed hungry on H.

During the process of training from J to L, "the fractional goal response  $r_{G_{ee}}$  first moves from L back toward J and then is evoked by J itself. Then when habit  $H_1IJ$  is formed this  $r_{G_{ee}}$ , now attached to J, becomes a part of J and is brought forward to path  $H_1$ . Thus a functional connection is established between the two related habit segments and becomes the basis of their subsequent unity" (5, p. 311). Here Hull assumes that  $r_{G_{ee}}$  is reinforced in the same way as any other response, that is to say, if its occurrence is rewarded, it tends again to be evoked. Confirmation is not necessary for its reinforcement, nor will disappointment cause its extinction. If Hull assumed the opposite,  $r_{G_{ee}}$  would not be brought forward to path H. If  $r_{G_{ee}}$  functioned like an anticipation of L and the large reward there, it would be frustrated if it occurred when the animal was running from H to J. For it would subsequently find J, not L, and a small reward instead of a large one. Thus  $r_{G_{ee}}$  would never be brought forward to path H because it would never be confirmed on this section, and if it occurred, it would suffer frustration which would presumably cause its extinction (5, p. 134). Therefore  $r_{G_{ee}}$  is functioning like an ordinary R in this part of Hull's explanation, and cannot be equated with anticipation in the way that Hull identifies it in a section entitled "The Realization of an Anticipation and its Frustration" (5, p. 133). Hence Hull is able to say that two sets of antedating goal reactions will occur at H and through their intermediary stimuli tend to evoke the first response leading to J. As he assumes that only one  $r_G$  will be conditioned to the sequence leading to N, the animal should go to J. The de-

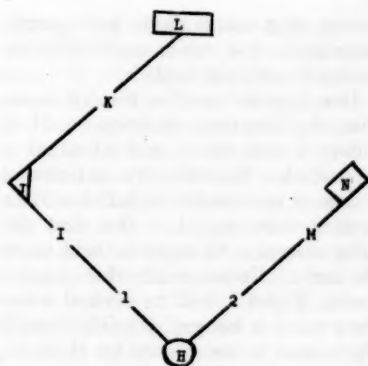


FIG. 3. A diagram of the maze in Hull's adaptation of Maier's behavioral insight experiment.

ciding factor Hull stresses is the disparity in the number of bonds existing between the stimulation arising at H and the two behavior segments. However, Hull again appears to ignore the fact that, during training, the  $r_{G_e}$  belonging to J and  $r_{G_{ee}}$  belonging to L would occur at H. It is to be remembered that the animal learned the two habits J to L and H to J before it was placed on H to learn to run towards N. Now  $r_{G_e}$  and  $r_{G_{ee}}$  ought to be evoked both because of the drive stimulus ( $S_D$ ) of hunger and because of the stimulation arriving at H. Thus three antedating responses should be connected to the habit H to N.

Hull can escape this consequence only if he treats the  $r_G$ s evoked at this point as expectations that would be falsified and subsequently extinguished on the animal's arrival at H. However, if he does this, he will not be able to use the argument which brings the anticipatory response belonging to L,  $r_{G_{ee}}$  down to H, according to the argument outlined above. It is difficult to see any other reason for Hull's assumption that the antedating responses occurring at H would not be conditioned to the sequence H to N

except that  $r_{GS}$  function as expectations and not as responses; Hull never makes this reason explicit.

It might be open to him to argue that the stimulus situation at H is different when the animal is trained to N, and that therefore the antedating reactions previously evoked would be evoked more weakly. But first the drive stimulus  $S_D$  must be held to be the same. Consequently the  $r_{GS}$  antedating J and L will be evoked when the animal is hungry, possibly even if the animal is not started on H at all. Second, however weakly these  $r_{GS}$  occur at the beginning of training to N at H, they will occur more strongly than at the beginning of training from H to J, and this path (H to N) will receive an *ex hypothesi* equal amount of reinforcement. Therefore, if anything, these  $r_{GS}$  should, after such a propitious start, be more strongly conditioned to responses leading to N.

These are no isolated examples such as might be due to an oversight. The antedating goal reaction has been treated as if it functioned as an expectation in other contexts. The inconvenient formal properties of an antedating goal reaction are suppressed when a prediction is required and "confirmation" is muddled with "reward." An example of this may be seen in an analysis of the double drive learning situation quoted with approval by Hull from Spence, Bergmann and Lippitt's discussion (8) of an experiment by Kendler (6).

Animals are trained while both thirsty and hungry to run a T maze with water in one arm of the T and food in the other. In the test trial they are made either hungry or thirsty. Kendler found, under these conditions, a tendency to choose appropriately. The result is difficult for Hull to explain.

If the animal turns right ( $R_R$ ) to that arm of the maze, it finds food and eats; if it turns left ( $R_L$ ) into that arm of the maze, it finds water and drinks. By analogy of the preceding analysis,  $S_{D_h}$  and  $S_{D_t}$  [the drive stimuli of hunger and thirst] will both be attached to  $R_R$  and  $R_L$ . So far as this mechanism alone goes, when the animal later enters the maze hungry but not thirsty the  $S_D$  alone will tend to evoke both  $R_R$  and  $R_L$  (5, p. 137).

Spence, Bergmann, and Lippitt attempt to give another analysis by introducing antedating goal reactions. "... the stimulus cues in the water arm and end box become, during training, conditioned to fractional anticipatory drinking acts ( $r_w$ ) and, in turn, the proprioceptive cues ( $s_w$ ) resulting from these anticipatory acts become conditioned to the response of entering this alley" (8, p. 549). A similar conditioning of anticipatory eating acts will take place in the food alley.

During the test series when only one drive is operative, the anticipatory act related to the goal for which the subject is motivated will, because of the greater strength of the particular drive stimulus, Hull's principle of stimulus dynamism, be much stronger (more vigorous) than the other, and hence will produce stronger proprioceptive cues. Thus, if the S is thirsty, proprioceptive cues from anticipatory drinking responses will be stronger than those from anticipatory eating. As these cues ( $s_w$ ) will be conditioned to the response of entering the alley leading to water, they will tend to give this response the greater excitatory strength (stimulus dynamism) (8, p. 549).

Now this explanation will not hold unless we treat anticipatory goal responses as expectations. In the above quotation it is assumed that the anticipatory goal responses of drinking and eating ( $r_w$  and  $r_t$ ) occur at the choice point simultaneously. Therefore  $s_w$  and  $s_t$  are produced also. Hence, as  $S_{D_h}$  and  $S_{D_t}$  in Hull's analysis quoted above, they should both be attached to the alternative responses. Consequently one of them alone should



evoke both the water and food turning tendencies. To say, as do Spence *et al.* (8, p. 549), that "the relative excitatory strengths of the two competing responses will be a function of the relative strengths of the two fractional anticipatory goal responses," is manifestly incorrect unless it is tacitly assumed that only confirmation and not reward leads to the conditioning of an anticipatory goal reaction. This assumption is not stated and does not seem to follow in any way from Hullian theory. In fact, Spence *et al.*, in an analysis in the very next paragraph of the latent learning experiment, explicitly make the opposite assumption. In this situation analyzed, an animal satiated for food and water is placed in a T maze and finds water in one arm and food in the other, and is then placed with its cage mates. Here

... the fractional anticipatory goal responses are much weaker, as the Ss do not actually eat or drink during the training period. The sight of food (or water) evokes conditioned fractional eating (drinking) responses. These, as described above, become conditioned to the cues of their respective alleys, *the reinforcement being provided by the social goal*, or possibly, as in higher order conditioning, by the sight of food or water itself (i.e., secondary reinforcement) [authors' italics].

Here the assumption is clearly stated that anticipatory goal responses are reinforced by goals which do not confirm them. It may be noted in passing that this analysis founders on the same objection as the last.

It has been put to the writer that there are additional premises which could save the deductions that have been considered, even though they could not stand as they are. The failure is attributed not to Hullian theory but to a lack of explicitness in the argument. The point that should be made explicit in the above derivations is that goal responses are incom-

patible. For Hull, learning requires both S-R contiguity and reinforcement. If the final stimulus ( $S_F$ ) is removed, so that  $R_F$  is blocked, some other incompatible response,  $R_X$ , must be made instead. This  $R_X$ , if reinforced, is carried forward as a fractional antedating  $r_X$  just as  $r_I$  was. The outcome of a competition between two incompatible reaction potentials is handled, in Hull's latest postulate set, by Corollary XIV, Postulate XIII. ("When the net reaction potentials ( $\bar{E}_R$ ) to two or more incompatible reactions (R) occur in an organism at the same instant, each in a magnitude greater than  $\bar{L}_R$ , only that reaction whose momentary reaction potential ( $\bar{E}_R$ ) is greatest will be evoked") (5, p. 13). Anticipatory responses are presumably incompatible; since only one can be made at a time, they cannot both be equally reinforced in the same end box.

Though it would greatly decrease the plausibility of the derivations of latent learning and reasoning if they hinged on the assumption that antedating responses were incompatible, such premises might preserve the logic of the Hullian deductions. Let us therefore examine what their application would do to the argument, taking as an instance the Kendler double-drive situation, on account of its simplicity. Let us first of all take the training situation. The two factors to watch are the strength of the conditioning of the stimuli at the choice point ( $S_c$ ) to the antedating goal responses ( $r_G$ ), and second, the strength of the conditioning of the fractional goal stimuli ( $s_G$ ) to the responses of turning to either goal; only these factors can influence the final choice of the animal in the test, all others being equal according to Hull. Let us suppose that the animal is trained by turning an equal number ( $x$ ) of trials

to each side, the trials to each side alternating in equal numbers. Let us recall that the animal is trained while both hungry and thirsty to run for a reward to food on the left of a T maze and to water on the right. Let us suppose that the strength of reinforcing factors on each side of the maze is balanced.

After a certain number of trials, one of the  $r_G$ s (let us say  $r_l$ ) will be evoked by  $S_e$ , the stimuli at the choice point, due to the fact that a set of trials to the food side occurred first in the training schedule. Suppose the number of trials in each set is  $x$ . Then  $r_l$  will be evoked by  $S_e$   $x$  trials before  $r_w$  could be. Assuming an increment in the tendency of  $S_e$  to evoke  $r_l$  each time its trace is reinforced to  $R_l$ ,  $r_l$  will have been strengthened  $x$  more times than  $r_w$ , its competitor, and therefore  $r_w$  could only be evoked by  $S_e$   $x$  trials later than  $r_l$  ( $x$  trials is the very earliest, but it would probably be only later, as once  $r_l$  is evoked by  $S_e$ , other reinforcing factors begin to favor it). Then  $r_l$  will be evoked by  $S_e$  at all events until the end of a set of trials to food, as the other  $r_G$ ,  $r_w$ , cannot take its place because  $S_e \rightarrow r_w$  is not increasing. We have already shown that  $r_w$  cannot take the place of  $r_l$  at  $S_e$  until at least  $x$  trials after  $r_l$  have appeared at  $S_e$ . But a set of trials is only  $x$  trials long. Hence  $r_w$  will not appear during the set of trials,  $x$  in number, when the animal is running to water. Even if  $r_l$  does not increase appreciably while the animal is running to water during  $x$  trials—as it ought, owing to reinforcement due to need reduction and secondary reward factors—it will at least hold the fort against  $r_w$ , which cannot appear till it is stronger than  $r_G$ . After the  $x$  trials, however, the siege is lifted, and  $r_G$  is enabled to get another  $x$  trials of increment. Thus  $r_w$  is

never able to supplant  $r_l$ , even if we set the factors of reinforcement of  $S_e \rightarrow r_l$  at the very lowest and make these vanishingly small.

Now if  $r_w$  is never evoked by  $S_e$ ,  $s_w$  cannot be reinforced to a response of turning left. Hence, though  $r_w$  would be aroused at  $S_e$  if the animal was made thirsty alone, the stimulation  $s_w$  and  $S_e$  present at the choice point could not determine the correct direction except by chance.

If  $r_l$  is aroused and is stronger, whatever the reason (for instance if the animal is made hungry), then the animal would most probably go to the water on the right. If an equal number of trials has been run to both sides,  $s_l$  will have been more frequently reinforced to  $R_{Rl}$  on the water side. This is because  $s_l$  will only appear and be reinforced when  $r_l$  has been evoked by  $S_e$ , and this can only be during a set of trials toward the food. The

probability is  $\frac{1}{x}$  that  $r_l$  will appear on the first trial of a set of  $x$  trials toward food, and thus, given that  $x$  is greater than 1, the chances are that  $s_l$  will be reinforced fewer times toward  $R_{Ll}$  (the food side) than toward  $R_{Rl}$  (the water side). The first set of trials toward water, after  $r_l$  has been evoked some time during the previous set toward food, will be a full one, and as the numbers of sets of trials towards each side are equal,  $s_l$  will be more frequently reinforced to  $R_{Rl}$  (the water side), on the average. The probability of this will increase with  $x$  (the number of trials in a set). It is assumed in calculating this probability that the chances of the appearance of  $r_l$  are equal for any one of the trials ( $x$  in a set) toward food.

The above argument showed that an assumption of incompatibility of antedating goal responses does not help the derivations if the number of

trials to each side is equal, and if they are made in equal batches. However, it might be said that an unfavorable training schedule has been chosen, and that a different type of training in which the numbers of trials are not kept rigorously equal, as between the sets of trials, could produce the desired learning in the animal. Let us therefore attempt to construct an optimum training schedule, allowing ourselves complete freedom. The trouble with the equal sets of trials was that  $r_w$  did not get evoked at the choice point at all. In order to circumvent this, we must give the animal more trials in the set toward the water after  $r_t$  has been evoked by  $S_c$ , so that  $r_w$  will be able to displace  $r_t$  at the choice point. In this way  $r_w$  will be evoked and will generate  $s_w$ , which will now be reinforced to the correct turn. But in doing this, we make the situation worse from another point of view. If we give more trials to water and continue until  $r_w$  is evoked by  $S_c$  instead of  $r_t$ ,  $s_t$  will be more often reinforced to the response of turning to water ( $R_{rt}$ ). Thus the tendency of  $s_t$  to be reinforced to  $R_{rt}$ , already present when the trials were precisely equated, will be aggravated. Thus if we wish  $S_c \rightarrow r_w \rightarrow s_w \rightarrow R_{rt}$  to take place at all we must inevitably strengthen  $S_c \rightarrow r_t \rightarrow s_t \rightarrow R_{rt}$  over  $S_c \rightarrow r_t \rightarrow s_t \rightarrow R_{Lt}$ . If we do not,  $s_w \rightarrow R_{rt}$  will never occur.

We are thus caught on the horns of a dilemma. It must also be clear that further training will not help. If we wish to strengthen  $s_t \rightarrow R_{Lt}$  we must first evoke  $S_c \rightarrow r_t$ . To do this we must reinforce  $s_w \rightarrow R_{Lt}$ , since when we run the animal to food  $r_w$  will first be evoked at  $S_c$ , as it has a larger reaction potential. It is stronger in virtue of the number of trials on which the trace of  $S_c$  was reinforced to  $R_w$  in the goal box, and any reinforcement which occurred when  $S_c$  evoked  $r_w$ , all

after  $r_w$  had drawn even with  $r_t$  and had displaced it at  $S_c$ . Now for  $r_t$  to appear at  $S_c$ , it must take at least the number of trials to the food side which occurred after  $S_c$  had evoked  $r_w$  when the animal was running to the water side. That means, however, that  $s_w$  will now be more strongly reinforced to  $R_{Lt}$  than to  $R_{rt}$  (water side). This will not give the correct prediction.

In the preceding discussion it has been assumed that the reinforcement of the trace of  $S_c$  to  $R_t$  or  $R_w$  is more effective than the reinforcement of  $S_c$  to  $r_t$  or  $r_w$ . In fact, the reinforcement of  $S_c$  to  $r_t$  or  $r_w$  has been set at a minimum. If, however, it is made larger than the reinforcement to the trace, the argument becomes somewhat simpler. Once an  $r_G$  is evoked and reinforced to the stimulus which evoked it, other  $r_G$ s simply cannot displace it. Each time the trace of this stimulus becomes reinforced to other responses, the bond of the stimulus itself to the  $r_G$  will always undergo a larger increment.

Another objection has been made against the preceding criticism. It has been said that the present writer assumes that in the examples discussed there is only one stimulus combination ( $S_c$ ) at the choice point that gets conditioned to responses on both sides. But actually the stimulus traces conditioned to the two  $r_G$ s are not the same. If food is on the right and water is on the left, then  $r_t$  and  $r_w$  are conditioned to different sets of place cues and response cues; viz., those on the right and left respectively. It is true that Hull did not make this distinction in his 1935 article. But in 1952, paraphrasing the Spence, Bergmann, and Lippitt analysis, he did distinguish clearly between "the stimulus traces of . . . looking right at the choice point x of the T ( $S_{xR}$ ) and beyond and the trace of  $S_{xL}$  (looking to

left at choice point x)" (5, p. 138). As a result, his ensuing derivation seems to avoid the difficulties brought up.

It is true that some stimulus traces will tend more strongly to evoke antedating eating responses, and others antedating drinking responses, in the Kendler type of situation. Once the animal has made the response which will make the appropriate, more strongly conditioned, stimuli impinge on its receptors, there is no longer a problem. Why should the animal choose to expose itself to one set of stimuli rather than the other? Looking right or looking left will exclude one set of stimuli. However, both sets of stimuli will tend, owing to previous learning, to evoke a response, though one more strongly than the other. If the choice is random, the animal will expose itself to either set of stimuli indifferently. Once one set of stimuli is made to appear in such a way that the other set will concomitantly disappear, the animal's behavior will be determined by it. Thus if the animal looks at the weaker set, it does not help to say that the other set would be stronger if the animal looked at it, for it is the set which is impinging which is at that moment stronger. It is the actual present stimuli that should determine the response, and not some other hypothetical ones, which would be stronger if only they were impinging. The stimuli from both sides which occur concomitantly will tend to arouse both  $r_{GS}$  equally, as they occurred together during training. These occur up to and at the choice point. The animal's behavior at the choice point cannot be governed by stimuli which it has not yet received. It is unlikely that Hull himself would have wished to take this line of defense. If we assume a differential exposure to stimuli from one

side, we do not need to have an explanation in terms of  $r_{GS}$ . The orthodox conditioning explanation which he rejects (5, p. 137) would do just as well.

Though the above examinations of lines of defense has produced nothing positive, it has at least clarified the issue. It is not sufficient for the appropriate antedating goal reaction to be evoked in the crucial trials; the fractional goal stimuli to which they give rise must also have been conditional to the correct response. What is needed is some mechanism which prevents the inappropriate antedating response from occurring, if the subsequent goal response is not like it. This can be ruled out. Alternately, there must be some mechanism which prevents the association of a fractional goal stimulus with a subsequent response, if the subsequent goal reaction is not the full version of the antedating reaction which gave rise to the fractional goal stimulus. However, such a mechanism could only work if assumptions of incompatibility of antedating goal responses were *not* made. Another way, and perhaps the most plausible, is to assume that only the correct antedating goal response is reinforced. Again it must be assumed that antedating responses are not incompatible. As these two last suggestions would involve closely similar assumptions, they will be discussed together.

An explicit introduction of these assumptions of confirmation into Hullian theory would raise considerable problems. And yet these must be faced. For if the above examination of the Hullian use of the anticipatory goal response is correct, many important derivations will not stand unless these assumptions are explicitly introduced. The problem may best be outlined by considering a concrete example. Sup-

pose a child expects some ice cream but is instead rewarded by being given a shilling; then if its expectation functions as a response, the next time in the same situation the child should expect ice cream even more strongly. Further, when it becomes hungry for ice cream it should tend to behave in such a way that it will be brought into the situation where, expecting ice cream, it received a shilling instead. In order to avoid such predictions, it must be stipulated that an anticipatory response will be reinforced only if it is followed by another response like itself. However, this other response must not be an anticipatory response because, as we have seen, anticipatory responses are evoked somewhat too frequently.

This means that there must for the animal be some way of distinguishing anticipatory responses from other responses. This separation of real responses from make-believe or anticipatory responses is also made necessary on another ground. Real responses must presumably be reinforced without the recurrence of another response like themselves.

Yet, if the anticipatory responses are in a different class from real or actual responses, in what does their similarity consist? Are they copies occurring in some special part of the musculature or in the central nervous system? Or are they muscular contractions carrying a special coding?

Further, granted that a satisfactory way of differentiating them is proposed, what leads to their initial differentiation when they are evoked? It has been supposed by Hull that an anticipatory goal response is evoked in identically the same way as any

other response. What, then, will determine, when a response is evoked, whether it will be classed as anticipatory or actual? This classification is necessary for the correct formula for reinforcement to be applied.

It seems, however, that even if the relevant premises were included in Hullian theory, and contradictory premises removed, its scientific status would be impaired.

As the system stands it appears that the deductions of insight, reasoning and most latent learning are not based upon the postulates but are logically fallacious. For an explanation of the phenomena avoiding these difficulties, see Deutsch (1, 2, 3).

#### REFERENCES

1. DEUTSCH, J. A. A new type of behaviour theory. *Brit. J. Psychol. (Gen. Section)*, 1953, **44**, 304-317.
2. DEUTSCH, J. A. A machine with insight. *Quart. J. exp. Psychol.*, 1954, **6**, 7-11.
3. DEUTSCH, J. A. A theory of insight, reasoning, and latent learning. *Brit. J. Psychol. (Gen. Section)*, 1956, **47**, 115-125.
4. HULL, C. L. The mechanism of the assembly of behavior segments on novel combinations suitable for problem solution. *Psychol. Rev.*, 1935, **42**, 219-245.
5. HULL, C. L. *A behavior system*. New Haven: Yale Univer. Press, 1952.
6. KENDLER, H. H. The influence of simultaneous hunger and thirst drives upon the learning of two opposed spatial responses of the white rat. *J. exp. Psychol.*, 1946, **36**, 212-220.
7. OSGOOD, C. E. *Method and theory in experimental psychology*. New York: Oxford Univer. Press, 1953.
8. SPENCE, K. W., BERGMANN, G., & LIPPITT, R. A study of simple learning under irrelevant motivational reward conditions. *J. exp. Psychol.*, 1950, **40**, 539-551.

(Received September 19, 1955)



## PERCEPTION: TOWARD THE RECOVERY OF A DEFINITION

JULIAN HOCHBERG

Cornell University

Perception, a word which once had a fairly limited meaning, has recently acquired a host of new (and frequently mutually contradictory [e.g., 26, 27]) ones. The shattering of the taxonomy of sensation, image, and percept (15, 30) made it difficult to maintain the older definitions (e.g., perception equals sensations plus images). Since then, usage has ranged from a narrowing of "perception" almost to the previous connotations of "sensation" (27), to a widening which includes almost all cognition—e.g., "Symbolization makes possible *perception at a distance* where the immediate percept is only a symbol of a distant event" (26). So loose has the term become that today Murray might call his battery the Thematic Perception Test, and not arouse any great comment. This diffuseness renders the term almost meaningless, and, even if there were no more serious consequences, it tends to promote a feeling of false unity and community of subject matter among diverse disciplines which really use the word in quite different fashions (cf. 29).

When the different definitions of the term are allowed to shift and interact in a single argument, however, serious confusion may result. When pre-, sub-, and unconscious perceptual processes are postulated (1, 17), when the results of presumably "perceptual" demonstrations are the occasion of profound epistemological conclusions (6, 12), and when questions are raised as to whether it is "really" perception (18, 21, 11) that is being investigated—at such points, it is not mere pedantry to ask questions about definitions. "Perception" frequently carries with it the

various connotations of "awareness," of a "discrimination" between stimuli, of a conviction of the "real" environmental presence of the perceived object, etc. These concepts are quite complex and as yet ill defined, and their casual intermixture may have methodological as well as theoretical consequences. For example, Bruner, Postman, and Rodrigues (4) attempted to demonstrate that the effect of the previously experienced normal color of an object on its presently perceived color increased as the ambiguity was increased by *increasing the time elapsed* between the presentation of the stimulus and the matching of that stimulus to a color wheel—a procedure more appropriate to the investigation of *memory* than to the traditional usage of *perception*. This is not to say that we must adhere to the old structuralist methods of slicing mental processes, but memorial effects in memory are not quite as critical a phenomenon as would be memorial effects in perception, and the distinction seems worth retaining when the existence of such effects is itself the question at hand. For most purposes such distinctions may actually be as unnecessary as the tendency to ignore them suggests, but this really depends upon the use to which the definition will be put.

It seems to me that there are at present at least two purposes for which a definition of perception is needed: (a) to delimit certain characteristics (phenomenal properties) of the *experience* of the subject, and (b) to distinguish between the immediate (primary) and less immediate (secondary) functional determinants we assume to be underlying the overt discriminative response.

Few attempts have been made, recently, at explicit definitions for either purpose, and this is what has generated the delightful Freudian paradoxes about "subceptual recognition," the ambiguity about motivational determinants in perception, and the confusion about "ambiguity."

#### *Programs for the Definition of Perception*

*Perceptual experience.* One aspect of classical consensual usage is given in a definition by MacLeod: Perception is that process by which things, events and relationships become phenomenally "here," "now" and "real." Some might object to the "privacy" and "subjective" nature of the construct; however, aside from such metaphysical objections, the actual problems of applying such definitions experimentally have largely been ignored. Since a "percept" is a construct, and not "directly measurable," its definition can be approached with different indices and at various levels.

At the most primitive level, a considerable gain in information could be achieved in many supposedly "perceptual" experiments (cf. 8) simply if subjects were asked whether their reports concerned objects "really seen" as present, rather than inferred or imagined. This is particularly important where the experimental findings are intended to reveal the fundamental dependence of perceptual experience *in general* upon some set of variables other than those immediately involved, e.g., dependence upon learning, upon motivation, etc. Thus, Hastorf (8) attempts to support the empiricist explanation of depth perception by showing that the known size of an object can determine its apparent distance: He instructs subjects that a rectangular stimulus, presented so that its visual distance is indeterminate, is

a calling card, an envelope, etc., and obtains different distance matches; the relevance of his findings would be considerably increased, however, by some assurance (in the face of some evidence to the contrary [10]) that his subjects' performance actually entailed perceiving the stimulus as instructed, and perceiving the distances as reported. Again, if we have attempted to demonstrate the effects of needs on the perception of length by rewarding longer lines (22), we require some assurance that the subjects are not merely reporting, without *seeing*, greater length (18). Where needed, greater certainty as to the degree to which "immediacy" characterizes the experience and as to the uni- or multidimensionality of the response system, could be achieved by the use of psychophysical and scaling procedures. *Whether or not a given response is perceptual in this sense is an empirical matter* (cf. Perky's demonstration of the continuity of image and percept [19]).

Since it seems unlikely that very many modern American experimental psychologists feel quite comfortable when brought face to face, in such bald fashion, with a construct of "conscious perceptual experience," the second purpose for which definition is required is probably more important at the moment (although it cannot be completely untied from this first problem).

*The temporal primacy of "perceptual" processes.* A recurring theme (by no means new) is the helpless dependence of perception on internal nonexteroceptive factors of motivation and past experience. Thus, our perceptions of the world—its spatial fabric (16), the objects within it (3), and their dimensions (13) and attributes (4)—are supposedly determined by desire and habit. A percept becomes merely a normal hallucination and, with all

knowledge based upon the senses, all our cognitive processes are contaminated and deluded. In extreme cases (6), solipsistic epistemologies—each man being viewed as in an independent, monadic universe produced by the projection of his trigonometric judgments (16) or of his unconscious desires (1) or memories (9)—have been offered as psychological “findings” with great energy and persuasiveness. We are not concerned here, however, with either the philosophical or the factual aspects of this matter, but rather with the definitions of perception involved, implicit or explicit.

By the old structuralist definition of perception, all of this is trivial and at least as old as Bishop Berkeley—by definition, a *percept* is *not* in correspondence with proximal stimulation, but is compounded of sensations and remembered images. *Sensations* alone would be in expected correspondence with the physical world of proximal stimulus energies, while the nature of the world of objects and space would by definition be determined by (in fact, composed of) the associated images of past experience and the judgment-like processes of *unbewussterschluss*. The Princeton demonstrations of purported experiential determinants in perception should occasion no great amazement to any one who has the traditional definition in mind, with its temporal primacy of sensation and associated memory-trace arousal *before* the arousal of the perceptual process. Similarly, there is little occasion for pleased surprise at the “perceptual defense” claims that unconscious recognition processes intervene between what, in classical terms, would be sensation (itself not always conscious [2]) and conscious perception or overt recognition. Another class of implicit definitions has evidently been under attack here, one which ignores

the old sensation-perception dichotomy (14).

However, if we wish to establish anything concerning the effects of, say, judgment as an independent variable upon perception as a dependent variable, we must be sure that the process we have observed as a dependent variable was not, itself, the judgmental one whose *effect* we were presumably investigating (11, 21). Postman (20) seems to tell us that the perceptual response—and, therefore, everything useful we can say about “perception”—follows the “laws of associative learning.” However, the main point has been glossed over by Postman, and tends to return us to the earlier problem of conscious experience: *not every behavior measured is an index of the perceptual nature of the response, and not all indices are equally good*. Let us look, therefore, at some examples of perceptual research in which implicit or explicit attempts were made to answer this question.

1. If we reward differently the two alternates of a simple reversible figure-ground pattern, the situation is vastly different from that of rewarding different line lengths. Schafer and Murphy (24) coupled monetary reward with one, monetary loss with the other of two sets of half-faces, which were then combined to form single reversible figure-ground units and presented for tachistoscopic recognition. Here, subjects who report recognition of the previously rewarded half-face cannot be “judging”: If we assume (23) that only one of the two faces can be figure at one time (there is some doubt about this, however: [27, 25]), and that the tachistoscopic exposure employed does not allow time for a reversal of figure and ground, the subject cannot correctly report a face other than the one he first sees, since he cannot (except by chance) know which it is. *The subject cannot*

*be giving a report which could equally well have been made in the absence of the stimulus.*

2. Wallach, O'Connell, and Neisser (28) found that the shadow of a static wire form, originally reported to be two-dimensional, appeared three-dimensional after viewing the shadow cast by the same wire form in rotation. This would constitute a clear demonstration of the effect of past experience on depth perception if we could be sure that this was not merely recollection by the subjects that "this pattern is one which was cast by a three-dimensional form." The experimenters therefore had the subjects continue to view the same static shadow pattern, and when reversals of perspective were reported with appropriate changes in relative sizes of the parts, concluded that they were dealing with perceived, rather than just judged, depth. The criterion of reversal has been used before (for example, by Hochberg and Hochberg [10] in testing the perceptual effects of "familiar size" on "represented" depth).

Why is perspective reversal considered evidence that the response is perceptual? (a) For one thing, spontaneous perspective reversal is known to occur in the presence of certain kinds of stimuli, and has not yet been reported to occur in cognitive processes *which take place in the absence of the specific stimulus*. (b) More important is the element of "additional information" already noted to be implicit in the Schafer-Murphy criterion: the second (reversed) description of the wire form is only one of an unlimited number of responses which the subject *could* make, but the one which he *does* make is a specific and appropriate function of the distal stimulus. As with figure-ground reversal, but in considerably more complex fashion, the patterns of (proximal) stimulation on the retina

must change their function in the delineation of the object (distal stimulus): what were seen as the opposite edges of the same surface become edges of two different ones, etc. The subject must abandon in its entirety his first description of the distal stimulus, yet the second report cannot be made independently of the *distal* stimulus, and depends upon its presence.

Even more striking are the related changes in length accompanying the three-dimensional form reversal. If some wire side *a* was previously seen as equal to and nearer than *b*, it now appears longer and farther away than *b*. This is to be expected in terms of the laws of perspective, but the speed with which this occurs, the ignorance most subjects display with respect to such laws, and their inability to compute such trigonometric relationships (5) make it very doubtful that the response has been a judgmental one.

To summarize, one set of implicit criteria of the perceptual nature of any given response appears to be that the experimenter has some reason to believe that the *presence of the stimulus, and its excitation of neural processes, are necessary (but perhaps not sufficient) for at least certain aspects of that response*. The experimenter should not, in principle, be able to predict the response completely without knowledge of the specific stimulus, else we are dealing neither with sensation nor with perception. To distinguish between the latter requires definition of another concept—that of *ambiguity*.

#### *Perception, Sensation, and Ambiguity*

Thus, we have a continuum in place of the older dichotomy between sensation and perception. At one extreme, complete psychophysical correspondence (7) obtains (including most of the area known as *sensation*), and our knowledge

of the variance of some aspect of the stimulus object is necessary and sufficient to predict completely the response variance of the subject; here, considerations of *needs* and *past experience* are at present gratuitous. At the other extreme, subjects' responses are completely independent of the presented stimulus, knowledge of which is neither necessary nor sufficient for prediction of response, and we are not dealing with perception at all, but with other psychological processes such as judgment, imagery, etc. Between these two extremes lies an important domain (corresponding to some extent to the classical area of *perception*) in which the stimulus accounts for some but not all of the response variance, and here it may be fruitful to inquire as to how much of the residual variance may be sought in factors of motivation and habit.

The criterion underlying this continuum is a concept of *ambiguity*, i.e., the degree of the inter- and intra-individual variability of relationship between stimulus and perceptual response we considered above. (Another definition of ambiguity might be attempted in terms of the extent to which different stimuli are perceived as having more or less of the quality of ambiguousness, but as little is known about this dimension as about the dimension of phenomenal *reality* or *presence* discussed above, and to which it seems closely related.) Such ambiguity can never be ascribed to a stimulus alone, but must depend also upon the choice of dimensions along which the subject is requested to make his response. Consider a visual stimulus under any specified constant conditions of seeing—say a red feather. Is it an ambiguous one? The answer depends in part upon the questions I ask the subject. If I ask psychophysical questions about the form,

color, size, etc., I can get very low response variability. If I ask about its warmth, its softness, or its charity, I suspect that response variability will increase, while with respect to flavor or sexual attractiveness, variability will reach really respectable levels. The Rorschach is not an ambiguous stimulus if I am seeking the answer, "ink-blot," but becomes so only if subjects are asked to respond along dimensions to which the stimulus is partly or wholly irrelevant. In short, whether we are dealing with "sensation," "perception," or "judgment" is not to be determined by either the stimulus or the response dimension alone, but by both as related by the task set.

#### SUMMARY

The concept of the *percept* or the *perceptual response*, if it is to retain any useful meaning, requires definition at least in those situations where contradictory or controversial statements are made about it. Two sets of defining operations are suggested: (a) psychophysical scaling of experimental situations in terms of the immediacy or perceptual quality of the experiences they arouse; (b) the requiring of responses which cannot be made, by the naive subject, in the complete absence of the stimulus. The degree to which the presented stimulus fails to determine the response marks the *ambiguity* of the psychophysical relationship under a given task set.

#### REFERENCES

1. BLUM, G. S. An experimental reunion of psychoanalytic theory with perceptual defence and vigilance. *J. abnorm. soc. Psychol.*, 1954, **49**, 94-98.
2. BORING, E. G. A history of introspection. *Psychol. Bull.*, 1953, **50**, 164-189.
3. BRUNER, J. S., & POSTMAN, L. Perception, cognition and behavior. *J. Pers.*, 1949, **18**, 14-31.



4. BRUNER, J. S., POSTMAN, L., & RODRIGUES, J. Expectation and the perception of color. *Amer. J. Psychol.*, 1951, **64**, 216-227.
5. BRUNSWIK, E. Statistical separation of perception, thinking, and attitudes. *Amer. Psychologist*, 1948, **3**, 342. (Abstract)
6. CANTRIL, H. *Understanding man's social behavior (preliminary notes)*. Princeton, N. J.: Office of Public Opinion Research, 1947.
7. GIBSON, J. J. *The perception of the visual world*. Boston: Houghton Mifflin, 1950.
8. HASTORF, A. H. The influence of suggestion on the relationship between stimulus size and perceived distance. *J. Psychol.*, 1950, **29**, 195-217.
9. HILGARD, E. R. The role of learning in perception. In R. R. Blake & G. V. Ramsey (Eds.), *Perception: an approach to personality*. New York: Ronald, 1951.
10. HOCHBERG, CAROL B., & HOCHBERG, J. E. Familiar size and the perception of depth. *J. Psychol.*, 1952, **34**, 341-345.
11. HOCHBERG, CAROL B., & HOCHBERG, J. E. Familiar size and subception in perceived depth. *J. Psychol.*, 1953, **36**, 341-345.
12. ITTLESON, W. H., & KILPATRICK, F. P. Perception. *Sci. Amer.*, 1951, **185**, 50-56.
13. ITTLESON, W. H. Size as a cue to distance: static localization. *Amer. J. Psychol.*, 1951, **64**, 54-67.
14. KOFFKA, K. *Principles of gestalt psychology*. New York: Harcourt, Brace, 1935.
15. KÜLPPE, O. Versuche über Abstraktion. *Ber. ü. d. I. Kongr. f. exp. Psychol.*, 1904, 56-68.
16. LAWRENCE, M. *Studies in human behavior*. Princeton, N. J.: Princeton Univ. Press, 1949.
17. MCCLEARY, R., & LAZARUS, R. S. Autonomic discrimination without awareness: an interim report. *J. Pers.*, 1949, **18**, 170-179.
18. PASTORE, N. Need as a determinant of perception. *J. Psychol.*, 1949, **28**, 457-475.
19. PERKY, C. W. An experimental study of imagination. *Amer. J. Psychol.*, 1910, **21**, 422-452.
20. POSTMAN, L. Experimental analysis of motivational factors in perception. In *Current theory and research in motivation*. Lincoln, Nebr.: Univ. of Nebraska Press, 1952.
21. PRATT, C. C. The role of past experience in visual perception. *J. Psychol.*, 1950, **30**, 85-107.
22. PROSHANSKY, H. M., & MURPHY, G. The effects of reward and punishment on perception. *J. Psychol.*, 1942, **13**, 295-305.
23. RUBIN, E. *Visuellwahrgenommene Figuren*. Copenhagen: Glydendalska Boghandel, 1921.
24. SCHAFER, R., & MURPHY, G. The role of autism in a visual figure-ground relationship. *J. exp. Psychol.*, 1943, **32**, 335-343.
25. SMITH, D. E. P., & HOCHBERG, J. E. The effect of "punishment" (electric shock) on figure-ground perception. *J. Psychol.*, 1954, **38**, 83-87.
26. SWEET, A. Some problems in the application of cognitive theory to personality functioning. *J. Pers.*, 1953, **23**, 41-51.
27. WALLACH, H. Some considerations concerning the relation between perception and cognition. *J. Pers.*, 1949, **18**, 6-13.
28. WALLACH, H., O'CONNELL, D. N., & NEISSER, U. The memory effect of visual perception of three-dimensional form. *J. exp. Psychol.*, 1953, **45**, 360-368.
29. WERNER, H. Introductory remarks. *J. Pers.*, 1949, **18**, 2-5.
30. WERTHEIMER, M. Experimentelle Studien über das Sehen von Bewegung. *Z. Psychol.*, 1912, **61**, 161-265.

(Received November 3, 1955)

## SECONDARY MOTIVATION THROUGH CLASSICAL CONDITIONING: A RECONSIDERATION OF THE NATURE OF BACKWARD CONDITIONING<sup>1</sup>

JOHN A. BARLOW

Denison University

Does backward conditioning occur? The answer involves reconsideration of what we mean by conditioning, and requires a brief review of several recent studies.

We have the authority of Pavlov, and of at least five more recent reports (3, 4, 10, 14, 15), that backward conditioning results in conditioned responses which are at best "insignificant and evanescent." Pavlov further reports that the backward CS eventually "becomes inhibitory" (9, p. 381).

If we define *conditioning as the process of modifying the effects of a stimulus on the behavior of an organism by associating it with a second stimulus* rather than in terms of the more conventional stimulus-substitution formulation, Pavlov's results *do show* backward conditioning. The backward CS does not acquire the properties of the UCS but *opposite properties*.

Spooner and Kellogg (13) present curves showing the conditioning process. They used 60 women students as subjects. A 1,000-cycle tone was the CS, a single make-break DC impulse of 0.2-second duration the UCS, and withdrawal of the right index finger the response to be conditioned. In this study a typical conditioning or learning curve was found for forward conditioning: few conditioned responses at first; then, as conditioning progressed, responses were made on every test trial. They

report that the backward conditioning curve "*is apparently the reverse of that in forward conditioning*" (italics in original [13, p. 334]). In this case a few conditioned responses were made during the first few test trials interspersed among the training trials; then, on further trials, conditioned responses stopped occurring.

Two problems arise in considering the results of Spooner and Kellogg. First, while their result tends to confirm the opposite trends in effect of the forward and backward CS, they were able to observe only occurrence or absence of finger withdrawal (Pavlov reports that salivation in his backward conditioning studies was actually inhibited, i.e., fell below normal). Secondly, there was no control for pseudo-conditioning. Fitzwater and Reisman (6), in a study similar to that of Spooner and Kellogg found no significant differences between the backward conditioning curve and the curve of responses for a pseudo-conditioning control group.

Barlow (1, 2) used 100 white rats as subjects, a 5-second change in illumination as the CS, and a single 10-second electric shock as the UCS. Twenty hours after conditioning, presence of the CS was made dependent upon the S's touching a bar now inserted through the conditioning apparatus. For half of the Ss (Subgroup 1) bar contact resulted in occurrence of the CS during test; for the other half (Subgroup 2) the CS was present at all times except during bar contact. After forward conditioning (onset of UCS coinciding with termination of CS) the CS acquired negative prop-

<sup>1</sup> This article is based upon part of a dissertation submitted to the Graduate School of Arts and Sciences, Duke University, in partial fulfillment of the requirements of the Ph.D. degree.

erties such that mean time touching the bar was greater for Subgroup 2 than Subgroup 1 ( $p < .01$ ). After backward conditioning (termination of UCS coinciding with onset of CS), the CS acquired positive properties. In this case, mean time of touching the bar was greater for Subgroup 1 than for Subgroup 2 during the test period ( $p < .01$ ). The differences in mean time spent touching the bar for control subgroups were insignificant. Thus the motivational properties acquired by the CS in this study were opposite after backward conditioning from those obtaining after forward conditioning.

Prior to Barlow's study, Coppock (5) reported that a stimulus associated with shock termination acquired positive properties. There is a large literature (see review by Miller, [8]) on the secondary reinforcing properties of a stimulus associated with food (therefore with reduction in hunger?). However, perhaps due to the fact that presenting the CS after electric shock is a backward conditioning procedure, Coppock's study is the first experiment showing the secondary properties for the subject of a stimulus which has been associated with shock reduction.

Coppock found that after white rats had received 90 5-second alternating-current shocks through the tail, immediately followed by a 1-second "blinking in illumination," the properties of this stimulus for the Ss were definitely modified. As a test, Coppock made occurrence of the blinking illumination contingent upon head movements to the previously nonpreferred side.

In the last test period, the group in which the beginning of the light had coincided with the end of shock had greater total duration of head movements toward the reinforced side than the groups in which the beginning of the light had preceded or followed shock termination ( $p = .02$  and  $.01$ , respectively) (5, p. 277).

Since the above studies, Smith and Buchanan have reported that "Rats trained with a black alley following shock made fewer errors and took fewer trials to reach a criterion in a black-white choice situation with black positive than rats trained with a white alley following shock" [ $p < .01$ ] (12, p. 126). Goodwin and Brownstein shocked rats in a black box, then let them escape into a white box. As a test, the Ss in one of the groups were put in a discrimination situation where an S was

"... simultaneously exposed to a box similar to the escape chamber and a neutral box. ... The shift in the number of choices of the box similar to the escape chamber between a series of preference trials given prior to training and the same box following training was significant at the .01 level of confidence" (7, p. 385).

#### CONCLUSION

From these studies we conclude, not that backward conditioning does not occur (or rarely occurs) or that "*forward and backward conditioning are fundamentally distinct processes*" (italics in original, [13, p. 334]); but that the backward CS later serves as a "signal" (to use Pavlov's term) just as does a forward CS. However it "signifies" something totally different: the termination instead of the onset of the UCS.

In the traditional conditioning procedure the response-to-be-conditioned is a response regularly elicited, initially, by the UCS, whether the UCS is shock, food, acid, or some other agent. If the CS precedes the UCS, this CS after conditioning signifies the onset of the UCS and elicits a CR which is similar to the UR observed to the UCS (and acquires motivational and/or reinforcing properties similar to those of the UCS). When the CS follows the UCS in training, a similar learning process results. The CS in this case signifies the termination of the UCS, and the

effect upon behavior of future presentation of this CS is similar to the effect of termination of the UCS. The acquired properties of forward and backward CSs are opposite, just as are the properties of the onset vs. the termination of any UCS with which they might respectively be associated.

## REFERENCES

1. BARLOW, J. A. Secondary motivation through classical conditioning. *Amer. Psychologist*, 1952, 7, 273. (Abstract)
2. BARLOW, J. A. Secondary motivational properties of a stimulus after backward vs. forward conditioning. *J. exp. Psychol.*, in press.
3. BERNSTEIN, A. L. Temporal factors in the formation of conditioned eyelid reactions in human subjects. *J. gen. Psychol.*, 1934, 10, 173-197.
4. CASON, H. Backward conditioned eyelid reactions. *J. exp. Psychol.*, 1935, 18, 599-611.
5. COPPOCK, H. W. Secondary reinforcing effect of a stimulus repeatedly presented after electric shock. *Amer. Psychologist*, 1951, 6, 277. (Abstract)
6. FITZWATER, M. E., & REISMAN, M. Comparison of forward, simultaneous, backward, and pseudo-conditioning. *J. exp. Psychol.*, 1952, 44, 211-214.
7. GOODSON, F. E., & BROWNSTEIN, A. Secondary reinforcing and motivating properties of stimuli contiguous with shock onset and termination. *J. comp. physiol. Psychol.*, 1955, 48, 381-386.
8. MILLER, N. E. Learnable drives and rewards. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951. Pp. 435-472.
9. PAVLOV, I. P. *Lectures on conditioned reflexes*. (Trans. by W. H. Gantt and G. Volborth.) New York: International Publ., 1928.
10. PORTER, J. M., JR. Backward conditioning of the eyelid response. *J. exp. Psychol.*, 1938, 23, 403-410.
11. RAZRAN, G. Backward conditioning. *Psychol. Bull.*, 1956, 53, 55-69.
12. SMITH, M. P., & BUCHANAN, G. Acquisition of secondary reward by cues associated with shock reduction. *J. exp. Psychol.*, 1954, 48, 123-126.
13. SPOONER, A., & KELLOGG, W. N. The backward conditioning curve. *Amer. J. Psychol.*, 1947, 60, 321-334.
14. WOLFLE, HELEN M. Time factors in conditioning finger-withdrawal. *J. gen. Psychol.*, 1930, 4, 372-378.
15. WOLFLE, HELEN M. Conditioning as a function of the interval between the conditioned and the original stimulus. *J. gen. Psychol.*, 1932, 7, 80-103.

(Received January 9, 1956)





# THE BRITISH JOURNAL OF PSYCHOLOGY

Edited by JAMES DREVER

Vol. XLVII. Part 3 November 1956 20s. net—\$3.50

**HUNICE BELBIN.** The effects of propaganda on recall, recognition and behaviour. I. The relationship between the different measures of propaganda effectiveness.

**JOYCE A. M. BROOKS and PHILIP E. VERNON.** A study of children's interests and comprehension at a science museum.

**HANNAH STEINBERG.** 'Abnormal behaviour' induced by nitrous oxide.

**HANS H. TOCH and WILLIAM H. ITTELSON.** The role of past experience in apparent movement: a reevaluation.

**J. A. DEUTSCH.** The statistical theory of signal after-effects and acuity.

**J. G. TAYLOR and S. PAPERT.** A theory of perceptual constancy.

## PUBLICATIONS RECENTLY RECEIVED

The subscription price per volume, payable in advance, is 50s. (or \$10) net (post free).

*Subscriptions may be sent to any bookseller or to the*

**CAMBRIDGE UNIVERSITY PRESS**

Bentley House, Euston Road, London, N.W. 1.